



UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA
Instituto Universitario de Microelectrónica Aplicada
Sistemas de información y Comunicaciones

Máster en Tecnologías de Telecomunicación



Trabajo Fin de Máster

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución en la Reserva Natural de las Dunas de Maspalomas

Autor: Judit Celia Viera Santana
Tutor(es): Gustavo Marrero Callicó
F. Javier Marcello Ruiz
Fecha: Septiembre 2016



t +34 928 451 086 | iuma@iuma.ulpgc.es
f +34 928 451 083 | www.iuma.ulpgc.es

Campus Universitario de Tafira
35017 Las Palmas de Gran Canaria



Máster en Tecnologías de Telecomunicación



Trabajo Fin de Máster

**Evaluación de técnicas avanzadas de clasificación
para la generación de cartografía de alta resolución
en la Reserva Natural de las Dunas de Maspalomas**

HOJA DE FIRMAS

Alumno/a:	Judit Celia Viera Santana	Fdo.:
Tutor/a:	Gustavo Marrero Callicó	Fdo.:
Tutor/a:	F. Javier Marcello Ruiz	Fdo.:

Fecha: Septiembre 2016





Máster en Tecnologías de Telecomunicación



Trabajo Fin de Máster

**Evaluación de técnicas avanzadas de clasificación
para la generación de cartografía de alta resolución
en la Reserva Natural de las Dunas de Maspalomas**

HOJA DE EVALUACIÓN

Calificación:

Presidente

Fdo.:

Secretario

Fdo.:

Vocal

Fdo.:

Fecha: Septiembre 2016



ÍNDICE

Capítulo 1: Introducción.....	3
1.1. Antecedentes.....	3
1.2. Objetivos.....	4
1.3. Estructura de la memoria.....	6
Capítulo 2: Técnicas de pre-procesado de imágenes WorldView 2.....	9
2.1. Fundamentos de la Teledetección.....	9
2.2. Estado del arte para los satélites de alta resolución: WorldView 2.....	14
2.3. Estado del arte en el pre-procesado de los datos de teledetección.....	17
2.3.1. Corrección de las distorsiones geométricas.....	18
2.3.2. Corrección de las distorsiones radiométricas.....	22
2.4. Resumen.....	25
Capítulo 3: Técnicas de procesamiento de imágenes de satélite.....	27
3.1. Metodología de Clasificación.....	28
3.2. Clasificación Supervisada.....	28
3.2.1. Fase de entrenamiento.....	29
3.2.1.1. Selección de la región de interés.....	30
3.2.1.2. Estudio de la separabilidad de las clases.....	30
3.2.2. Fase de asignación.....	32
3.2.2.1. Máxima probabilidad.....	33
3.2.2.2. Máquina de Vectores Soporte (SVM).....	35
3.2.3. Fase de Post-procesado.....	52
3.3. Evaluación de calidad de la clasificación.....	52
3.4. Técnica <i>Linear Spectral Unmixing</i> (LSU).....	54
3.5. Resumen.....	55
Capítulo 4: Generación de cartografía de vegetación y sustrato en Maspalomas.....	57
4.1. Datos y clases de interés.....	57
4.1.1. Imagen WorldView 2.....	57
4.1.2. Clases de interés.....	58
4.1.3. Medidas de campo.....	61
4.1.4. Cartografías de interés de Maspalomas.....	64

4.2. Metodología.....	67
4.3. Software ENVI	70
4.4. Resultados para la imagen sin fusionar	71
4.4.1. Regiones de interés (ROIs) y cálculo de la separabilidad	71
4.4.1.1. ROI de entrenamiento para la imagen original	71
4.4.1.2. ROI de evaluación para la imagen original	76
4.4.2. Generación y Aplicación de la Máscara.....	78
4.4.3 Clasificación supervisada y matrices de confusión.	79
4.4.4. Parametrización y evaluación del SVM.....	84
4.4.5. <i>Linear Spectral Unmixing</i>	92
4.4.6. Filtrado <i>Majority</i> sobre la imagen original	96
4.4.7. Análisis de resultados obtenidos sobre la imagen original.....	99
4.5. Resultados para la imagen fusionada (GS).....	102
4.5.1. Regiones de interés (ROIs) y cálculo de la separabilidad	103
4.5.1.1. ROI de entrenamiento para la imagen fusionada	103
4.5.1.2. ROI de evaluación para la imagen fusionada	105
4.5.2. Generación y aplicación de la Máscara.....	107
4.5.3. Clasificación supervisada y matrices de confusión	107
4.5.3.1. Clasificaciones empleando <i>Maximum Likelihood</i> (ML)	107
4.5.3.2. Clasificación empleando <i>Support Vector Machine</i>	109
4.5.4. Filtrado <i>Majority</i> sobre la imagen fusionada.....	112
4.5.4.1. <i>Majority</i> sobre los resultados de <i>Maximum Likelihood</i>	112
4.5.4.2. <i>Majority</i> sobre los resultados de <i>Support Vector Machine</i>	113
4.5.5. Análisis de resultados sobre la imagen fusionada	114
4.6. Comparativa de resultados obtenidos	115
4.7. Resumen	118
Capítulo 5: Revisión de objetivos y resultados	119
5.1. Revisión de objetivos	119
5.2. Revisión de resultados	120
5.3. Líneas futuras de trabajo.....	121
BIBLIOGRAFÍA	123

PLIEGO DE CONDICIONES.....	135
PRESUPUESTO	139
ANEXOS	145
Anexo 1	147
Anexo 2	155
Anexo 3	179

Agradecimientos

A mis tutores Javier Marcello y Gustavo Marrero por la dedicación, el apoyo, trabajo y esfuerzo ofrecidos para que este Trabajo Fin de Máster fuera posible.

Al Doctor Antonio Ignacio Hernández Cordero por su labor de apoyo a la hora de evaluar las cartografías obtenidas. Por enseñarme un mundo que desconocía por completo. Aún recuerdo su empeño al intentar que nosotros tuviésemos la capacidad de distinguir el “*Juncus acutus*” por su forma de erizo.

Al proyecto ARTeMISat (CGL2013-46674-R), financiado por el Ministerio de Economía y Competitividad de España, por suministrar la imagen WorldView 2 y los datos de campo empleados en este Trabajo Fin de Máster.

A los profesores que han marcado mi carrera enseñándome lo que sé y lo que soy.

A mi familia, especialmente mis padres y mi hermano, por sus grandes dosis de paciencia y cariño. Por motivarme, apoyarme y comprender mis silencios.

A mis amigos, en especial a mi niña Nuria por animarme una y mil veces sin importar cual fuese el problema.

MEMORIA

Capítulo 1: Introducción

1.1. Antecedentes

Desde sus inicios, la teledetección ha proporcionado una fuente de datos muy valiosa para el estudio de las superficies terrestres a diferentes escalas espaciales y temporales [1]. Se hace muy complicado establecer un punto de inicio en la historia de la teledetección dado que los primeros sistemas fueron militares. Las primeras reseñas nos llegan de 1960, año en que Estados Unidos pone en órbita el primer satélite de observación de la Tierra. Con el paso de los años se han ido desarrollando, entre otros, satélites de muy alta resolución espacial, sensores hiperespectrales (cientos de bandas espectrales) y satélites en el dominio de las microondas.

La finalidad de la teledetección en el espectro óptico-IR es identificar y caracterizar los materiales de la superficie terrestre, así como los procesos que en ella ocurren, a partir de la radiación electromagnética procedente de la propia superficie terrestre, que se obtiene como la reflejada de la radiación que le llega del Sol.

Con la teledetección se abren numerosos campos de aplicación y estudio como pueden ser: el análisis oceanográfico en cuanto a salinidad, turbidez, corrientes, temperatura, localización de pesquerías, predicciones de deshielo, perfiles atmosféricos, seguimiento de la vegetación, cartografía, determinación de la deforestación, de incendios forestales y sus daños, etc. En la actualidad se dispone de un gran número de satélites orbitando alrededor de La Tierra. Esto ha sido posible gracias a la colaboración de distintos países y empresas públicas y privadas que han permitido, con sus lanzamientos, disponer de una extensa base de datos de imágenes. Además, para la aplicación de interés en este proyecto, contamos con satélites que nos proporcionan imágenes de alta resolución como, por ejemplo, Ikonos, Geoeye y WorldView 2.

Por otro lado, la teledetección también abarca todo el trabajo de procesamiento e interpretación realizado a posteriori con las imágenes captadas [2]. La información recogida en dichas imágenes no se haya exenta de errores. Dichos errores son de una naturaleza muy variada, pudiéndose agrupar en radiométricos y geométricos. Los errores radiométricos son aquellos que modifican el valor real de la energía captada por el sensor del satélite. Entre las causas de los errores radiométricos se encuentran

aquellas debidas a la presencia de la atmósfera, la orografía del terreno, mal funcionamiento del sensor, fallos de transmisión o recepción de las imágenes, etc. Por otro lado, los errores geométricos modifican la geometría de la imagen y son debidos a la curvatura de la Tierra, a su propio movimiento giratorio, distorsiones provocadas por el sensor, etc.

Una vez eliminados los errores de la imagen, se dispone de los medios necesarios para realizar, por ejemplo, la clasificación eficiente para discriminar las distintas coberturas que conforman la imagen. La clasificación supone, en numerosos casos, la fase culminante del análisis de imágenes ya que clasificar una imagen implica categorizarla, de tal forma que el nivel digital (ND) de un píxel determina la clase donde ha de ser incluido [3]. De esta forma, si un píxel satisface unos determinados criterios es asignado a una clase que se corresponde con esos criterios.

En este contexto, este Trabajo Fin de Máster se centra en la clasificación de una zona de estudio para obtener, así, cartografías de vegetación y sustrato de especial interés. Para realizar la clasificación supervisada, se sigue el siguiente esquema:

- **Fase de entrenamiento.** Consiste en definir el rango de ND o valores de reflectividad que identifica a cada categoría.
- **Fase de asignación.** Agrupación de los píxeles de la imagen en cada una de esas categorías tras la aplicación del algoritmo específico.
- **Verificación de resultados.** Las técnicas de verificación son aplicables tanto al tratamiento digital como al análisis visual y su objetivo es evaluar la fiabilidad de la cartografía generada.

1.2. Objetivos

El objetivo principal de este trabajo será la realización de un estudio comparativo de técnicas avanzadas de clasificación que permitan obtener como resultado una cartografía de vegetación y sustrato de la Reserva Natural de las Dunas de Maspalomas (Gran Canaria). Para ello emplearemos una imagen de muy alta resolución proporcionada por el satélite multispectral WorldView 2. Adicionalmente, se dispone de una base de datos proveniente del estudio de campo realizado utilizando un espectroradiómetro simultáneamente con el paso del satélite. A partir de la evaluación de los resultados obtenidos, debemos ser capaces de identificar el procedimiento y la técnica más robusta para la clasificación.

A continuación, detallaremos cada uno de los objetivos planteados al comienzo de este trabajo:

- ✓ Debemos estudiar el satélite que vamos a emplear en este proyecto: WorldView 2. Recopilaremos y estudiaremos la documentación necesaria para conocer los aspectos técnicos, el funcionamiento del satélite y, especialmente, los niveles de procesamiento y el formato de los datos, ya que nos proporcionará la imagen de estudio para este proyecto.
- ✓ Estudiaremos los pre-procesados necesarios para garantizar que los datos de la imagen se corresponden adecuadamente con la reflectividad emitida por la superficie terrestre. Es decir, analizaremos las distintas fuentes de error y sus correspondientes correcciones, además de técnicas como la fusión de imágenes para mejorar la resolución espacial de las bandas multispectrales.
- ✓ Debemos hacer un estudio teórico de las técnicas de clasificación, así como las medidas de calidad y evaluación de las mismas.
- ✓ Para desarrollar las tareas posteriores, es necesario tener un conocimiento previo del software a emplear: la herramienta Environment for Visualizing Images (ENVI).
- ✓ Se estudiará y ejecutará la fase de entrenamiento. Es una de las fases más complejas de este proyecto ya que en ella trataremos de definir e identificar con rigor cada una de las categorías que pretendemos discriminar en la imagen, teniendo en cuenta su propia variabilidad en la zona de estudio. En este caso diferenciaremos, tanto sustratos (arena seca, arena húmeda, costra salina, charca, callaos y aluviales) como tipos de vegetación ("*Tamarix canariensis*", "*Traganum moquinii*", "*Juncus acutus*", "*Suaeda mollis*", "*Tetraena fontanesii*" y "*Launaea arborescens*").
- ✓ Realizaremos la clasificación de imágenes de alta resolución empleando ENVI para la obtención de una cartografía de la vegetación y el sustrato de la Reserva Natural de las Dunas de Maspalomas, siendo ésta una de las tareas de mayor importancia a lo largo del proyecto. En ella aplicaremos diferentes clasificadores para tratar de asignar cada uno de los píxeles de la imagen a una de las clases previamente seleccionadas. Adicionalmente, estudiaremos de forma especial el clasificador *Support Vector Machine*.
- ✓ Una vez realizada la clasificación, debemos evaluar y analizar los resultados obtenidos para obtener información respecto al grado de acierto y la robustez de

los métodos empleados. Para ello, cuantificaremos el grado de acierto comparando dicha clasificación con la información real del terreno. Será necesaria la generación de una base de datos de evaluación, seleccionando en la imagen otras regiones distintas a las de entrenamiento, correspondientes a cada clase.

- ✓ Intentaremos optimizar los resultados obtenidos aplicando técnicas como la fusión de imágenes.
- ✓ Suavizaremos los resultados finales obtenidos para una mejora visual utilizando la técnica de filtrado denominada *Majority*.

En definitiva, el principal objetivo será la identificación del método más robusto para la clasificación de vegetación y sustratos en la Reserva Natural de las Dunas de Maspalomas.

1.3. Estructura de la memoria

Este trabajo se ha estructurado en seis capítulos y dos anexos que a continuación desglosaremos someramente.

El primer capítulo es en el que nos encontramos e incluye una breve introducción y los objetivos de este estudio.

El segundo describe los fundamentos de la teledetección y los sensores de alta resolución así como con una serie de características generales y espectrales del Satélite WorldView 2. Además, analizamos las técnicas de pre-procesado de nuestras imágenes satelitales.

El tercer capítulo presenta de forma teórica cada uno de los pasos a seguir en el proceso de clasificación: entrenamiento, clasificación supervisada, post-clasificación y evaluación.

En el cuarto capítulo se detallarán los resultados obtenidos en la generación de la cartografía de vegetación y sustrato a las Dunas de Maspalomas, de forma que nos permitan, además de seleccionar el método de clasificación más robusto, realizar un análisis comparativo. Igualmente, observaremos qué ocurre con los resultados si aplicamos, por ejemplo, una metodología de fusión de imágenes.

El quinto capítulo es el de conclusiones generales. En él analizaremos el grado de cumplimiento de los objetivos iniciales y extraeremos conclusiones a partir de los

resultados del capítulo anterior. Además, plantearemos diversas líneas futuras de trabajo dentro de este ámbito.

Además de las referencias bibliográficas, pliego de condiciones y presupuesto, hemos añadido una serie de anexos. En cuanto a los anexos, en primer lugar tenemos el anexo 1 compuesto por los pares de separabilidad obtenidos para cada una de las ROIs generadas a lo largo del trabajo. Asimismo se han incluido en el anexo 2 las matrices de confusión correspondientes a las evaluaciones realizadas durante el estudio de nuestra imagen. Por último, incluimos en el anexo 3 el código generado en Matlab para obtener la cartografía de interés con la paleta de colores adecuada para cada una de las clases.

Capítulo 2: Técnicas de pre-procesado de imágenes WorldView 2

Este capítulo se centra tanto en el estudio de la teledetección como en el conocimiento y análisis del satélite a emplear. Dicho satélite es el WorldView 2. Además, se describirán brevemente las técnicas de pre-procesado de los datos de teledetección.

2.1. Fundamentos de la Teledetección

La teledetección es una técnica para obtener información acerca de objetos sin que los instrumentos empleados para adquirir los datos estén en contacto directo con el objeto [4].

En teledetección hay tres elementos esenciales: fuente de energía, objetivo o escena y sensor [5]. Brevemente, explicaremos cada uno de estos elementos:

- **La fuente de energía** es la que “ilumina” el objetivo emitiendo una onda electromagnética. También puede medir el calor que se desprende de la superficie del objetivo (infrarrojo térmico). En nuestro caso la fuente de energía es el Sol.
- **El objetivo o escena** es la porción de la superficie terrestre espacial observada por el satélite. Su dimensión varía, en función de la resolución del sensor, de unos pocos m^2 a algunos miles de km^2 .
- **El sensor** mide la energía solar (radiación electromagnética) reflejada por el objetivo. El sensor puede encontrarse en un satélite o en un avión, sobrevolando el objetivo a una altura de pocos centenares de metros hasta distancias de 36000 kilómetros en el caso de los satélites meteorológicos.

Por ejemplo, una superficie blanca refleja cantidades iguales de radiación de todas las longitudes de onda de luz visible, mientras que una hoja verde refleja menos radiación en las partes rojas y azules del espectro que en la parte verde. Ello ofrece un exceso de luz verde (en comparación con la azul y la roja), por lo que la hoja parece verde. La composición de la reflexión electromagnética para cada longitud de onda (la

firma espectral) nos ofrece información sobre la superficie que emite o refleja la radiación.

La capacidad de los satélites de distinguir entre varias firmas espectrales es vital para su uso en cartografía, donde es esencial la distinción entre los distintos tipos de coberturas terrestres [6]. En la Figura 2.1 podemos apreciar, a modo de ejemplo, la firma espectral de la vegetación y de distintos tipos de aguas y suelos.

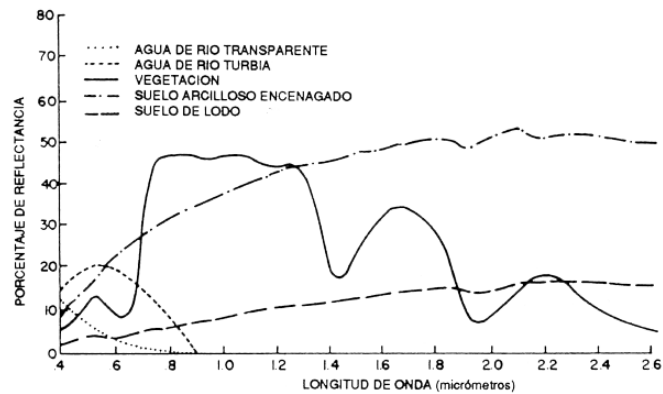


Figura 2.1. Firma espectral de diversas características naturales de la superficie de la Tierra [7].

El ojo humano sólo puede percibir radiaciones de una gama limitada del espectro electromagnético (Figura 2.2). Por eso, los instrumentos para teledetección, fuera de las longitudes de onda visibles, representan realmente una extensión de nuestro campo visual y proporcionan acceso a información adicional sobre el mundo físico que nos rodea [8].

La radiación electromagnética de una superficie es una reflexión (luz reflejada) o una emisión (radiación emitida desde la superficie). La luz solar reflejada sólo puede ser medida, obviamente, durante el día, mientras que la emisión puede medirse en cualquier momento.

La temperatura de la superficie es un factor clave de la emisión. El Sol tiene una temperatura superficial de 6.000 grados Kelvin (K) y una emisión máxima de la gama de luz visible. Una superficie con una temperatura de unos 1.000 K, por ejemplo un incendio en el Amazonas, tiene su emisión máxima en el espectro infrarrojo medio [9]. La temperatura superficial de la Tierra es de unos 290 K y tiene una emisión máxima a unos 14 micrómetros, también llamada banda térmica de infrarrojos. Existe una correlación directa entre la temperatura superficial y el grado de emisiones de una longitud de onda determinada. La temperatura superficial puede calcularse sobre la base de teledetección de emisiones térmicas de infrarrojos.

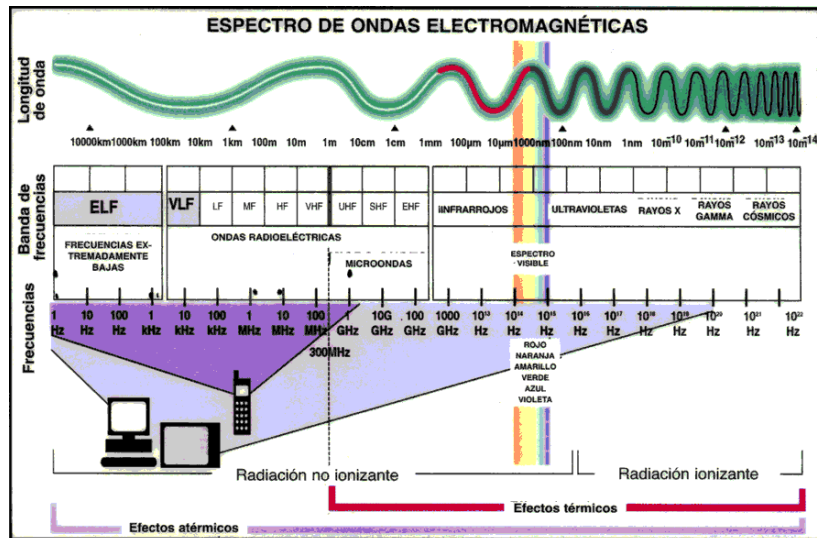


Figura 2.2. Espectro Electromagnético [10].

Otro factor a tener en cuenta es la interferencia atmosférica. La radiación del Sol y la reflexión de la superficie de la Tierra atraviesan la atmósfera antes de llegar al sensor del satélite. El contenido atmosférico de los gases absorbe parte de la radiación terrestre. El ozono actúa como una barrera casi completa a la radiación ultravioleta, y la mayoría de radiación en la gama de 9.5 – 10 micrómetros queda absorbida. El vapor de agua y el dióxido de carbono son gases que absorben la radiación en muchas longitudes de onda distintas.

Las longitudes de onda donde la mayoría de la radiación puede atravesar la atmósfera se llaman “ventanas atmosféricas” (Figura 2.3). Afortunadamente, también puede pasar una gran cantidad de luz visible (de lo contrario, la luz del día nunca llegaría a la superficie terrestre). La atmósfera también es casi 100% transparente para ciertas gamas del espectro cercano al infrarrojo, lo que posibilita la observación por satélite con un mínimo de distorsión atmosférica. La gama térmica infrarroja de 10 -12 micrómetros se usa para medir las temperaturas superficiales del suelo, el agua y las nubes. Además, existe otra ventana en la banda de las microondas.

Incluso, si la teledetección se realiza en ventanas atmosféricas, ésta se ve interferida hasta cierto punto por la dispersión y la absorción de la atmósfera. Por eso, la energía captada por el sensor de teledetección queda ligeramente distorsionada y debe ajustarse mediante un posterior procesamiento digital de la imagen que efectúe la corrección atmosférica.

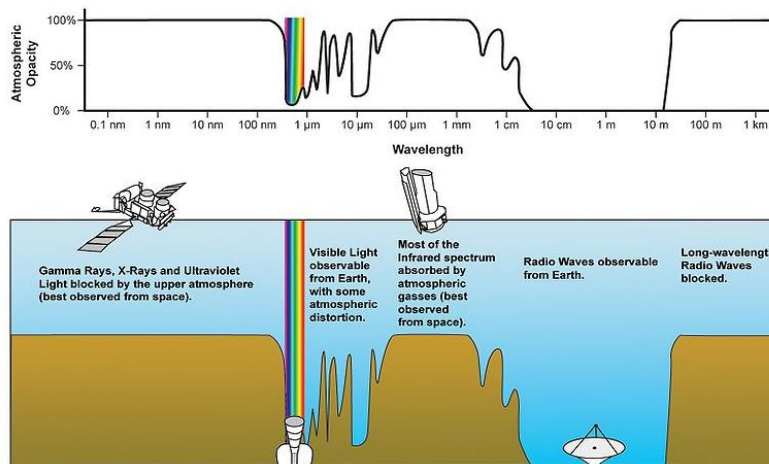


Figura 2.3. Absorción atmosférica para diferentes longitudes de onda [11].

Por último, deberíamos analizar los sensores de alta resolución puesto que los sistemas de percepción remota difieren en la resolución, es decir, el nivel de detalle con que pueden capturar las imágenes, su frecuencia temporal, etc. Desde este punto de vista podemos considerar los diferentes tipos de resolución que se detallan a continuación. Antes, debemos tener en cuenta que la radiación (emitida o reflejada) de la superficie terrestre es un fenómeno continuo en 4 dimensiones (espacio, tiempo, longitud de onda y radiancia). Un sensor debe muestrear en este continuo discretizándolo [12]. El modo en que esta discretización se lleva a cabo define los cuatro tipos de resolución con los que se trabaja en teledetección:

- **Resolución espacial:** es el tamaño del píxel, es decir, designa al objeto más pequeño que puede ser distinguido sobre una imagen. En algunos casos se emplea el concepto de IFOV (campo instantáneo de visión) que se define como la sección angular (en radianes) observada en un momento determinado. La relación entre IFOV y el tamaño del píxel (d) puede verse en la ecuación 2.1 y en la figura 2.4

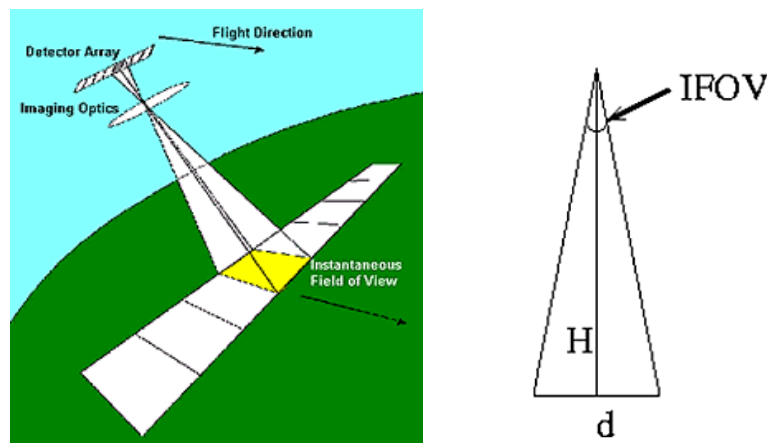


Figura 2.4. Relación IFOV- tamaño de píxel [13].

$$d = 2H \tan\left(\frac{IFOV}{2}\right) \quad (2.1)$$

donde d es el tamaño de píxel y H la distancia del sensor a la superficie terrestre. La resolución de estos sensores óptico-electrónicos depende de la altura orbital, la velocidad de exploración y el número de detectores.

La resolución espacial tiene un papel muy importante en la interpretación de la imagen porque marca el detalle que se ofrece. Además está relacionada con la escala de trabajo y con la fiabilidad de la interpretación.

- **Resolución espectral:** indica el número y la anchura de las bandas espectrales que puede discriminar el sensor, es decir, en las que el sensor puede captar radiación electromagnética. En principio, cuantas más bandas incluya un sensor mejor, ya que cada banda constituye una variable para caracterizar la superficie captada. Por otro lado, es preferible que estas bandas sean estrechas para que su poder discriminante se incremente. Si las bandas son muy anchas van a recoger valores promediados que ocultarán elementos de diferenciación. Por ejemplo resulta mucho más informativo contar con 3 bandas (rojo, azul y verde) en el visible que con una sola banda (imagen en escala de gris). El número de bandas y su localización en el espectro va a depender de los objetivos que se pretendan cubrir con la puesta en funcionamiento del sensor.
- **Resolución radiométrica:** indica la sensibilidad del sensor, es decir la capacidad de discriminar entre pequeñas variaciones en la radiación que capta. Suele expresarse mediante el número de bits necesarios que se precisan para almacenar cada píxel. Por ejemplo si un sensor utiliza 8 bits, proporciona $2^8 = 256$ niveles de energía (Niveles Digitales, ND) que pueden ser captados. Cuanto mayor sea la precisión radiométrica mayor número de detalles podrán captarse en la imagen.
- **Resolución temporal:** indica el intervalo de tiempo entre cada imagen obtenida por la plataforma espacial en la misma zona. Es de pocos minutos en el caso de los satélites geoestacionarios y variable en el caso de los satélites heliosíncronos de órbita baja. Recientemente se han desarrollado satélites de muy alta resolución espacial con sensores orientables, de manera que la zona de la que captan imágenes es relativamente independiente de la posición orbital que ocupe el satélite con lo que la resolución temporal pasa a ser algo más flexible.

Los diferentes tipos de resolución están muy relacionados. Generalmente, a mayor resolución espacial disminuye la temporal y también la espectral. El problema es que cada aumento de resolución multiplica el volumen de datos que el sensor debe preprocesar y enviar a las estaciones de Tierra. Sin embargo, como se ha indicado, la aparición de sensores de orientación modificable ha permitido compatibilizar alta resolución espacial y alta resolución temporal a costa de una toma de imágenes sistemática.

El diseño de los sensores de un satélite se hace teniendo en cuenta todas las consideraciones hechas hasta ahora. Salvo excepciones, se va a trabajar en las ventanas atmosféricas y se va a seleccionar aquella combinación de regiones del espectro que mayor información van a proporcionar acerca de los fenómenos que se quieren estudiar.

2.2. Estado del arte para los satélites de alta resolución: WorldView 2

WorldView 2 pertenece a la compañía americana DigitalGlobe y es el primer satélite comercial de muy alta resolución espacial provisto de 8 bandas multiespectrales. Se lanzó el 8 de octubre de 2009 a bordo de un cohete Delta 7920 desde la base de la fuerza aérea estadounidense en Vandenberg.

Se ubicó en una órbita polar heliosíncrona que se encuentra a una altitud de 770 kms, proporcionando, en el nadir, un ancho de exploración de 16.4 km y una resolución espacial de 46 cm para el canal pancromático y de 1.85 m para el multiespectral. Su periodo orbital es de aproximadamente 100 minutos y el tiempo promedio de revisita depende de la latitud y varía entre 1 y 4 días. En consecuencia, es capaz de captar datos diariamente de una superficie de casi 975.000 km^2 .

Está previsto que WorldView 2 esté operativo hasta el año 2017. En la Figura 2.5 se muestran los diferentes subsistemas que configuran la plataforma WorldView 2.

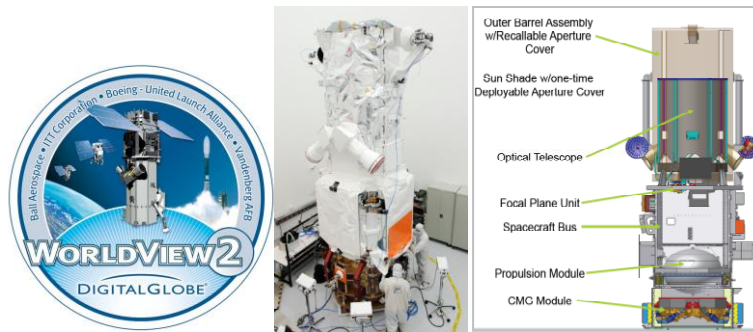


Figura 2.5. Subsistemas del satélite WorldView 2 [14].

Las principales ventajas frente a otros satélites de teledetección de alta resolución son las siguientes:

- **Mayor resolución espectral**

Worldview2 es el primer satélite comercial en proporcionar alta resolución espacial e imágenes multispectrales de 8 bandas. Así, además de las cuatro bandas multispectrales típicas (azul, verde, rojo e IR cercano), incluye 4 bandas adicionales para mejorar el análisis multispectral y permitir nuevas aplicaciones.

- **Mayor agilidad**

La serie de satélites WorldView son las primeras plataformas comerciales con capacidad para controlar los momentos de fuerza generados por los giroscopios. Esta tecnología de alto rendimiento proporciona una aceleración hasta 10 veces más que la de otros actuadores de control de actitud y mejora la maniobrabilidad y la capacidad de orientación. El tiempo de giro se reduce de más de 60 segundos a sólo 9 segundos para cubrir 300 km, permitiendo obtener imágenes de diferentes zonas en un pase único orbital.

- **Mayor Capacidad y Revisita**

Gracias a su mayor agilidad, Worldview2 puede obtener imágenes multispectrales de diferentes áreas en un solo pase. Worldview2 tiene una capacidad de exploración de hasta 975.000 kilómetros cuadrados por día. La combinación de una mayor agilidad en el Worldview2 y la altura orbital le permite tener una revisita casi diaria.

- **Mejor precisión**

La tecnología avanzada de posicionamiento de Worldview2 es lo que permite mejoras significativas en su precisión de geolocalización. La especificación de precisión se ha mejorado hasta 6.5 m CE90 sin ningún tratamiento adicional, uso de modelo de

elevación, ni puntos de control en tierra, pudiendo llegar hasta 2 m con datos adicionales.

En cuanto a las características espectrales de Worldview2 (Figura 2.6), éste lleva un instrumento que genera una imagen pancromática de alta resolución espacial (0.46 m en el nadir) y ocho bandas espectrales de resolución espacial cuatro veces peor (1.85 m en el nadir).

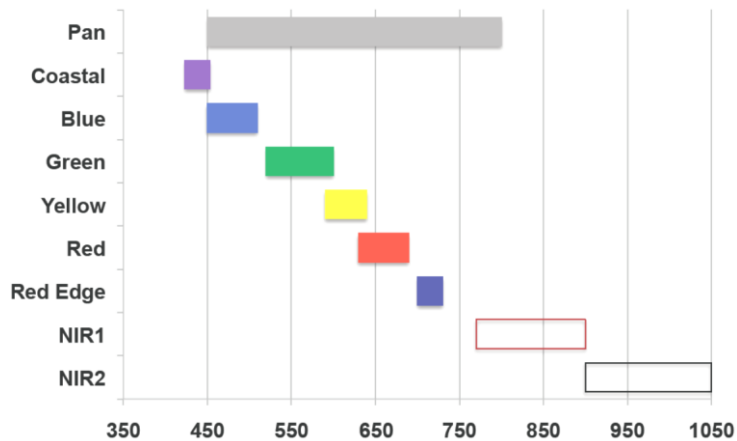


Figura 265. Bandas espectrales del satélite WorldView2 [14].

Las bandas multispectrales son capaces de proporcionar una precisión de color excelente, permitiendo el desarrollo de nuevas aplicaciones. Como se indicó, además de las cuatro bandas multispectrales tradicionales en el azul, verde, rojo e infrarrojo cercano, se incluyen cuatro bandas adicionales: una centrada en longitudes de onda más corta que el azul, aproximadamente en 427 nm; una banda amarilla, a 608 nm; una banda en el borde del rojo, centrada estratégicamente en, aproximadamente, 724 nm al ser el inicio de la parte de alta reflectividad de la respuesta de la vegetación, y una adicional en el infrarrojo cercano, pero a mayor longitud de onda, centrada aproximadamente en 949 nm, que es sensible al vapor de agua atmosférico.

Las principales características de las nuevas bandas incluidas son:

- **Azul costero** (400-450 nm): Ayuda en la realización de análisis vegetal. Mayor penetración en el agua por lo que ayuda en estudios marino-costeros y del agua. Muy útil en los estudios batimétricos. Tiene el potencial para mejorar las técnicas de corrección atmosférica.
- **Amarillo** (585-625 nm): Muy importante para la clasificación. Detecta la "amarillez" particular de la vegetación, tanto en tierra como en el agua.

- **Borde rojo** (705-745 nm): Muy valiosa para medir la salud de plantas y ayudar en la clasificación de la vegetación. Está más enfocado a la detección de la luz roja por la clorofila en vegetaciones sanas y es muy útil en la clasificación de suelos desnudos, carreteras, caminos y coberturas geológicas.
- **Infrarrojo cercano - NIR2** (860-1040 nm): Permite el análisis de la vegetación y estudios de la biomasa. Ayuda para la separación de cuerpos de agua y vegetación así como en la identificación de tipos de vegetación y en la separación entre tipos de suelos.

Es de interés saber que las imágenes Worldview2 están disponibles para su adquisición según 3 niveles de procesamiento [6]:

- *Basic*, que incluyen únicamente la corrección radiométrica.
- *Standard/Ortho-ready*, que incluyen la corrección radiométrica y geométrica básica (sin ortorectificación).
- *Stereo*, que incluyen 2 escenas superpuestas con ángulo de visión adecuado.

Las bandas multiespectrales y la pancromática se pueden adquirir de forma independiente o conjuntamente. En general, es necesario solicitar un polígono con una superficie en km² cuyo valor mínimo es variable según se trate de una imagen de archivo o programada. Estos productos se entregan con una alta resolución radiométrica de 16 bits y si el usuario lo requiere, la imagen puede entregarse con una resolución radiométrica de 8 bits ya sea en formato GeoTIFF o NITF 2.0 o 2.1.

2.3. Estado del arte en el pre-procesado de los datos de teledetección

El propósito de este apartado es la discusión de la naturaleza de los errores geométricos y radiométricos en imágenes de teledetección que nos permita encontrar procedimientos de desarrollo computacional para usarlos en su compensación.

Se pueden considerar las siguientes fuentes de error en la imagen:

- Producida por los instrumentos
- Producida por la atmósfera
- Producida por el terreno

Los errores se pueden clasificar de una manera genérica como: errores geométricos y radiométricos. Cuando los datos de las imágenes son registrados por sensores

embarcados en satélites, éstos pueden contener errores en geometría y en la medida de los valores de brillo (ND) de los píxeles. Este último efecto provoca los denominados errores radiométricos y pueden ser consecuencia de la instrumentación usada al registrar los datos, de la dependencia de la longitud de onda en la radiación solar y del efecto atmosférico. Los errores geométricos en la imagen se pueden deber a diversas causas: errores sistemáticos y no sistemáticos, la curvatura de la Tierra y variaciones incontroladas en posición y actitud de la plataforma del sensor.

Hay una serie de condicionantes externos a los sensores que afectan en las medidas de las superficies que adquieren los distintos sistemas. El grado de influencia puede ser mayor o menor dependiendo de la plataforma de la que procedan los datos. Vamos a pasar a analizar brevemente las distintas distorsiones y sus posibles correcciones.

2.3.1. Corrección de las distorsiones geométricas

Con el nombre de Corrección Geométrica se engloban las técnicas de corrección de las distorsiones que producen las principales fuentes de error en el momento de la adquisición de la imagen, como son el cabeceo (*pitch*), alabeo (*roll*), giro lateral (*yaw*) y las provocadas por la trayectoria, altura y velocidad de la plataforma que sustenta al sensor. A éstas hay que añadir los efectos del relieve y de la esfericidad terrestre. Así mismo, se incluyen en éstas técnicas las de georreferenciación, es decir, las de asignar una posición geográfica a los píxeles de nuestra imagen.

Los principales factores o fuentes de error geométrico en la imagen son:

- Rotación de la Tierra
- Ancho del campo de visión de algunos sensores
- Curvatura de la Tierra
- Variaciones de altitud, actitud y velocidad de la plataforma
- Efecto panorámico
- Relieve del terreno

Estas fuentes de error van a producir unas distorsiones desde el punto de vista geométrico que pueden ser tratadas en la mayoría de los casos conjuntamente, es decir, aplicando una corrección se pueden minimizar los errores procedentes de distintas fuentes.

Dichos errores se deben corregir empleando métodos paramétricos y no paramétricos. En los métodos paramétricos se introducen de forma directa y, mediante

sistemas e instrumentación de posicionamiento, los parámetros que intervienen. Por el contrario, en los métodos no paramétricos los parámetros de transformación se deducen a partir de puntos de control.

Para realizar la corrección geométrica se pueden seguir dos caminos:

- **Remuestreo de la imagen.**

Una vez que tenemos la imagen en su “situación” corregida, hay que transferir el nivel digital empleando para ello el de la imagen original. Se hace habitualmente de tres formas:

- Vecino más próximo: en la celda de la imagen corregida el ND corresponde al del píxel más cercano de la imagen original.
- Interpolación bi-lineal: promedia los ND de los cuatro píxeles más cercanos en el original. Difumina ligeramente los contrastes espaciales de la imagen original
- Convolución cúbica: considera los ND de los dieciséis píxeles más próximos. Requiere un volumen de cálculo elevado y el promediado es mayor.

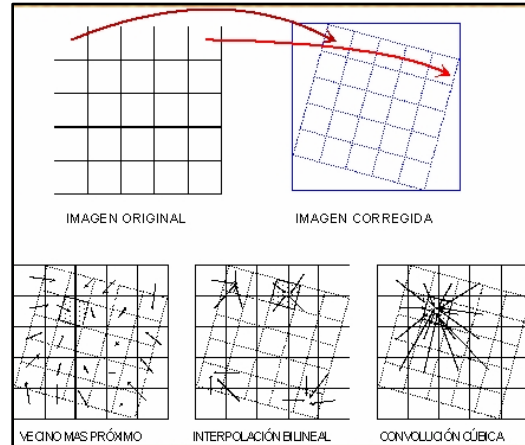


Figura 2.7. Remuestreo de la imagen [15].

- **Fusión de imágenes.**

Las técnicas de fusión de datos en teledetección se desarrollan debido al condicionante de diseño de los sensores de adquisición de imágenes, ya que gracias a éste se produce una relación inversa entre la resolución espectral de los mismos y la resolución espacial a la que pueden registrar datos. Por este motivo se han desarrollado metodologías cuyo objetivo último es transformar las imágenes multispectrales

incorporando el detalle espacial a partir de canales pancromáticos de alta resolución. Para hacer referencia a este concepto se emplean habitualmente términos "correlación multiresolución", "integración multifuente" o "fusión de datos", entre otros. Nosotros nos referiremos a estas técnicas como fusión de imágenes a nivel de píxel (*pansharpening*). En la Figura 2.8 se muestra un ejemplo de la técnica de fusión de imágenes que venimos desarrollando. En primer lugar tendríamos nuestra imagen pancromática (Figura 2.8 (a)) que fusionada a la imagen multispectral de la Figura 2.8 (b) nos daría como resultado la imagen fusionada de la Figura 2.8 (c).



Figura 2.8. (a) Imagen pancromática, (b) imagen multispectral y (c) imagen fusionada [16].

En la actualidad se dispone de un gran número de algoritmos de fusión [17], [18], [19], [20], [21] principalmente aplicados para la fusión de imágenes ópticas e IR cercanas.

En sus inicios, las técnicas más populares fueron las basadas en operaciones aritméticas, destacando los algoritmos de Brovey, *Synthetic Variable Ratio* o *Ratio Enhancement*, y las basadas en la sustitución de bandas tras la aplicación de un transformada, destacando el Análisis de Componentes Principales (PCA, *Principal Component Analysis*), la transformada Intensidad-Brillo-Saturación (IHS, *Intensity-Hue-Saturation*) o el algoritmo Gram-Schmidt (GS). La utilización de estos algoritmos está muy extendida dada la baja complejidad computacional que presentan. Sin embargo, algunas de ellas proporcionan imágenes fusionadas de menor calidad espectral, es decir cuyo color presenta distorsiones respecto al color de las imágenes multispectrales originales. Esto impide su uso en diferentes aplicaciones en el área de la teledetección, como son la clasificación de imágenes o la detección de cambios. Más recientemente, para el tratamiento de datos de satélite de nuevos sensores con mayor

número de bandas, como es el caso de Worldview-2 o de los sensores hiperespectrales, se ha desarrollado nuevos algoritmos como, por ejemplo, el *Hyperspectral Colour Sharpening* (HCS) [22].

Para solventar las limitaciones espectrales de los algoritmos mencionados, surgieron técnicas que inyectan la información de alta frecuencia, destacando los métodos HPF (*High-Pass-Filtering*), HPM (*High-Pass-Modulation*) o el basado en la aplicación de filtros paso alto en el dominio de Fourier [23]. Sin embargo, los métodos que utilizan el análisis multiresolución, y fundamentalmente la Transformada Wavelet Discreta (TWD), son los más populares para disminuir la distorsión espectral. En particular, para lograr resultados óptimos de fusión, diversos esquemas basados en wavelets han sido propuestos por varios investigadores [24], destacando los algoritmos de Mallat y *À trous*, cuya principal diferencia se refiere al sentido en el que se realiza la estrategia multiresolución, pues en el primer caso se diezma la imagen mientras que para *À trous* no se aplica ningún diezmo. Posteriormente, otras transformadas como Curvelets, Ridgelets o Contourlets se han propuesto como alternativas más eficientes, desde el punto de vista de la representación de la información, que la TWD [25], [26]. Estas nuevas transformadas permiten una extracción eficaz del detalle espacial en diferentes direcciones pero son computacionalmente muy complejas.

Otra estrategia seguida ha sido también la combinación de diferentes esquemas de fusión que puedan lograr una mejor calidad de los resultados, intentando aprovechar los beneficios de cada técnica individual.

En este trabajo se empleará la técnica de fusión de imágenes basada en la metodología de Gram-Schmidt. Este método se basa en un algoritmo general de ortogonalización de vectores, la ortogonalización de Gram-Schmidt.

En el método GS [27], la resolución espacial de la imagen MS se mejora mediante la fusión de la imagen de alta resolución Pan (pancromática) con las bandas de baja resolución espacial de la MS (multiespectral). De acuerdo a la descripción de los autores, las principales etapas del método se citan a continuación (Figura 2.9):

1. Una imagen pancromática de menor resolución espacial es simulada (**S**).
2. La transformación de Gram-Schmidt se realiza en la simulación realizada de menor resolución espacial de la imagen pancromática y bandas multiespectrales de menor resolución espacial (**MS**). La imagen

pancromática de menor resolución espacial simulada se emplea como la primera banda en la transformación de Gram-Schmidt (GS_1).

3. Las estadísticas de la imagen pancromática de resolución espacial más alta se ajustan para que coincida con las estadísticas de la primera banda transformada (GS_1) resultante de la transformación de Gram-Schmidt para producir una imagen pancromática de mayor resolución espacial modificada (\tilde{P}).
4. La imagen pancromática de mayor resolución espacial modificada (\tilde{P}) sustituye a la primera banda transformada (GS_1) resultante de la transformación de Gram-Schmidt para producir un nuevo conjunto de bandas transformadas.
5. La transformación inversa de Gram-Schmidt se lleva a cabo en el nuevo conjunto de bandas transformadas para producir la imagen MS de resolución espacial mejorada.

$$\begin{pmatrix} S \\ MS_1 \\ \vdots \\ MS_N \end{pmatrix} \xrightarrow[\text{transform}]{GS} \begin{pmatrix} GS_1 \\ GS_2 \\ \vdots \\ GS_{N+1} \end{pmatrix} \xrightarrow[\text{with modified P}]{\text{replace } GS_1} \begin{pmatrix} \tilde{P} \\ GS_2 \\ \vdots \\ GS_{N+1} \end{pmatrix} \xrightarrow[\text{transform}]{\text{Inverse GS}} \begin{pmatrix} X \\ F_1 \\ \vdots \\ F_N \end{pmatrix}$$

Figura 2.9. Método pan-sharpening de Gram-Schmidt.

2.3.2. Corrección de las distorsiones radiométricas

Los mecanismos que afectan e intervienen en la medida de los valores de brillo de los píxeles pueden ser fuentes de distorsión radiométrica. En primer lugar, la distribución del brillo en la imagen para cada una de las bandas dadas puede ser diferente en la escena y en el terreno. En segundo lugar el ND relativo de un píxel singular de una banda a otra puede estar distorsionado comparado con la reflectividad espectral correspondiente a la región en el terreno. Ambos tipos de errores pueden ser el resultado conjunto de la presencia de la atmósfera y de la transmisión del medio en relación con la radiación que llega al sensor, y puede ser el resultado también de efectos en la instrumentación.

En contraste con la corrección geométrica, en donde todas las fuentes de error suelen ser rectificadas conjuntamente, los procesos de corrección radiométrica tienen que ser específicos para la naturaleza de la distorsión.

Las técnicas más habituales de corrección radiométrica sin tener en cuenta el efecto de la atmósfera son:

- i. Corrección radiométrica del bandeo, en sensores de barrido producido por desajustes en el calibrado de los detectores.
- ii. Corrección radiométrica de líneas o píxeles perdidos, producido por fallos del sensor o de transmisión de los datos.
- iii. Corrección o Normalización radiométrica horaria.
- iv. Corrección radiométrica del efecto por columnas.

Las imágenes suministradas ya incorporan las anteriores correcciones. Por ese motivo no se detallarán.

- **Corrección atmosférica**

Entre el sensor y la superficie terrestre se interpone la atmósfera, que interactúa de formas diversas con el flujo radiante. La atmósfera está compuesta de distintos gases que producen: (a) la absorción de la energía en determinadas bandas del espectro, (b) la dispersión (o *scattering*) de determinados flujos en función de su longitud de onda y (c) la emisión, inherente a todo cuerpo caliente. Todos estos procesos producen modificaciones en la radiación originalmente propagada entre la cubierta terrestre y el sensor [15], [28].

Para compensar los efectos atmosféricos, parámetros tales como la distribución de aerosoles, la cantidad de vapor de agua y visibilidad de la escena deben ser conocidos. Como las mediciones directas de estas propiedades de la atmósfera rara vez están disponibles, se han desarrollado diferentes estrategias. Habitualmente, la corrección atmosférica se basa en la utilización de imágenes con métodos como el *Dark Object Substraction* (DOS) [29], que corrige el efecto de dispersión aditivo; y Coseno del ángulo cenital solar (COST) [30], que también tiene en cuenta el efecto multiplicativo de transmisión; *QUick Atmospheric Correction* (QUAC) [31], que se basa en el hallazgo empírico de que la reflectancia promedio de un material de espectros no depende de la escena. El *Simplified Method for the Atmospheric Correction* (SMAC) [32] es un método semi-empírico diseñado para la corrección atmosférica de una gran serie de datos adquiridos por sensores de gran campo de visión. El *Empirical Line Calibration* (ELC) [33] utiliza la reflectancia de la superficie medida de campo de una serie de patrones de calibración invariantes en el tiempo para calcular la reflectancia

para cada banda. A continuación, se desarrolla una ecuación de regresión para cada banda de una imagen para realizar la corrección atmosférica. Otro modelo más complejo se aproxima a la atmósfera y, por lo general, utilizan un código de transferencia radiativa precisa (RTC) para corregir los efectos atmosféricos, tales como *MODerate resolution atmospheric TRANsmission* (MODTRAN) [34], [35], *Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes* (FLAASH) [36], *ATmospheric CORrection* (ATCOR) [37], [38] y *Second Simulation of a Satellite Signal in the Solar Spectrum* (6S) [39], [40].

En este trabajo, se emplea el algoritmo de corrección atmosférica basado en el modelo *Second Simulation of a Satellite Signal in the Solar Spectrum* (6S). El *Second Simulation of a Satellite Signal in the Solar Spectrum* (6S) es un modelo avanzado de transferencia radiativa diseñado para simular la reflexión de la radiación solar en condiciones de una atmósfera libre de nubes, según las condiciones específicas geométricas y espectrales. Este modelo tiene en cuenta los principales parámetros atmosféricos para modelar la dispersión y la absorción que produce la atmósfera en la longitud de onda del canal del satélite. El algoritmo 6S es utilizado para generar las LUTs (*Look-Up Tables*) en los algoritmos de corrección atmosférica del sensor Modis (*Moderate Resolution Imaging Spectroradiometer*) de la NASA [41], [42], [43].

El código fuente de este algoritmo es libre bajo licencia GNU y ha sido programado en lenguaje Fortran. Se trata de un programa de línea de comandos que obtiene de un fichero de entrada los parámetros de configuración necesarios para la generación de los valores de los parámetros atmosféricos modelados.

La configuración de entrada del modelo se divide en 5 partes principales: condiciones geométricas, modelado atmosférico, alturas del área de estudio y del satélite, condiciones espectrales y reflectancia del suelo.

Una vez configurados los archivos necesarios para la corrección de cada uno de los canales del satélite se ejecuta el modelo, obteniendo unos archivos de salida con los valores de los principales parámetros atmosféricos que intervienen en los fenómenos de absorción y *scattering* que produce la atmosfera. La expresión para la corrección atmosférica de las imágenes del satélite WorldView 2 propuesta, para el modelo 6s, es:

$$\rho_{\lambda} = \frac{y}{1+(x_c * y)} \quad (2.2)$$
$$y = (x_a * L_{\lambda}) - x_b$$

donde,

- ρ_λ es la reflectancia corregida atmosféricamente.
- L_λ es la radiancia medida por el satélite en [$watts/(m^2 \cdot st \cdot \mu m)$].
- X_a es la inversa de la transmitancia atmosférica.
- X_b es el scattering de la atmósfera.
- X_c es el albedo atmosférico para la luz isotrópica.

De esta forma se realiza una conversión de la radiancia medida por el satélite a reflectividad de la superficie terrestre, ya corregida atmosféricamente, mediante el uso de tres variables generadas por el modelo 6S, para las condiciones geométricas y atmosféricas indicadas por los archivos de configuración.

2.4. Resumen

En este capítulo hemos estudiado las técnicas de pre-procesado de imágenes de satélite, en especial del satélite WorldView 2 puesto que ha sido el encargado de proporcionar a este Trabajo Fin de Máster la imagen de estudio. Inicialmente hemos resumido los fundamentos de la teledetección, hemos analizado el espectro electromagnético, los efectos de la absorción atmosférica, etc. A continuación, se han detallado los aspectos importantes de los satélites de alta resolución y nos hemos centrado en el conocimiento y análisis del satélite WorldView 2 ya que, tal y como hemos dicho anteriormente, ha sido el encargado de proporcionar la imagen de estudio. Una vez conocidos los problemas con los que se encuentra el satélite para capturar la imagen, se han descrito las técnicas de pre-procesado aplicables a los datos de teledetección. Dichas técnicas serán principalmente la corrección radiométrica, geométrica y atmosférica. Se han descrito los procedimientos para llevar a cabo esas correcciones y contar con una imagen corregida para nuestro estudio.

Capítulo 3: Técnicas de procesamiento de imágenes de satélite

En este capítulo, explicaremos en detalle una de las técnicas de procesamiento de imágenes de satélite más empleada: la clasificación. Adicionalmente trataremos de una manera menos detallada la técnica *Linear Spectral Unmixing*, con la que realizaremos un pequeño estudio a lo largo de este Trabajo Fin de Máster.

Para la mayor parte de los usuarios de la teledetección, la clasificación supone la fase culminante del tratamiento digital de imágenes. Como fruto de la clasificación digital se obtiene una cartografía e inventario de las categorías objeto de estudio. La imagen multibanda se convierte en otra imagen, del mismo tamaño y características de la original, con la importante diferencia de que el nivel digital que define cada píxel no tiene relación con la radiancia detectada por el sensor, sino que se trata de una etiqueta que identifica la categoría asignada a ese píxel [44]. A partir de una imagen como ésta puede fácilmente generarse cartografía temática, así como un inventario estadístico del número de píxeles asignados a cada clase.

Conviene considerar que puede abordarse una clasificación exclusivamente a partir de los niveles digitales de la imagen, ya que las categorías temáticas suelen definirse de modo relativo a las condiciones específicas de la escena a clasificar [45]. En otras palabras, la clasificación digital no busca una definición absoluta de cada cubierta, que pudiera ser aplicable a cualquier imagen, sino más bien una caracterización relativa, válida para una determinada imagen y un territorio concreto. Con este planteamiento, no resulta preciso conocer detalladamente las condiciones de adquisición: basta identificar en la imagen las clases a discriminar, sin pretender que esa identificación sea extrapolable a otras situaciones. En resumen, no se trata de establecer un modelo entre medidas digitales y parámetros físicos, sino medidas digitales, para un sensor, una zona y una fecha determinada, y tipos de cubierta. En este sentido, la clasificación supone un enfoque más cualitativo, si bien, por tratarse de un proceso numérico, incorpora un apartado estadístico ciertamente importante. Lógicamente esto supone el nivel de interpretación más elemental.

Una clasificación digital basada exclusivamente sobre niveles digitales puede

denominarse propiamente clasificación espectral. Con ser muy rica esta información no siempre es capaz de separar cubiertas pues, puede haber distintas categorías temáticas con un comportamiento espectral similar [46]. En ese caso, se recomienda acudir a las otras dos dimensiones de información que contiene una imagen: la espacial y la temporal. En el primer caso, se trataría de incorporar a la clasificación el contexto espacial en el que aparece cada píxel. En el segundo, se aportarían datos sobre su evolución estacional, lo cual es especialmente crítico para la discriminación de cubiertas vegetales.

3.1. Metodología de Clasificación

En capítulos anteriores hemos detallado el pre-procesado que se realiza sobre los datos de teledetección. Una vez concluida esa etapa, debemos definir la metodología de clasificación, que consta de varias fases totalmente diferenciadas:

- **Fase de entrenamiento:** consiste en caracterizar los patrones que definen en la imagen a las distintas categorías objetivo.
- **Fase de asignación:** consiste en asignar cada uno de los píxeles de la imagen a una de las clases previamente definidas.
- **Fase de post-clasificación:** consiste en el empleo de una serie de filtros para mejorar los resultados obtenidos tras la clasificación.

3.2. Clasificación Supervisada

El punto más importante de nuestro Trabajo Fin de Máster es la clasificación, porque su objetivo principal es precisamente la generación de mapas temáticos de la Reserva Natural en estudio.

Existen dos tipos de clasificación, la clasificación supervisada y la no supervisada. Estas dos clasificaciones parten de conceptos muy diferentes por eso deben analizarse de forma individual. En este capítulo, centraremos la explicación en la clasificación supervisada puesto que es la empleada en el desarrollo de nuestro proyecto.

La clasificación supervisada parte de un cierto conocimiento de la zona de estudio, adquirido por experiencia previa o trabajo de campo. Esta mayor familiaridad con la zona de estudio permite al intérprete delimitar sobre la imagen unas áreas suficientemente representativas. En la Figura 3.1 podemos observar un esquema del

proceso seguido en la clasificación supervisada.

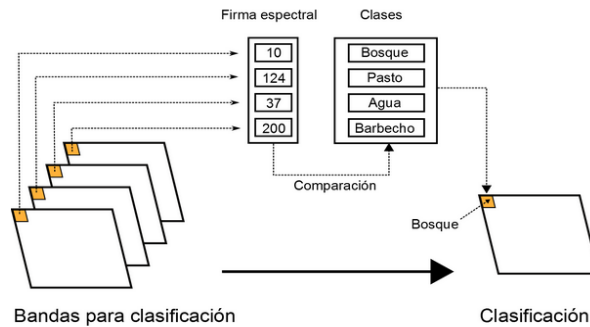


Figura 3.1. Esquema del proceso de clasificación supervisada [47].

Dentro de la clasificación supervisada, debemos tratar dos puntos fundamentales:

- Fase de entrenamiento
- Fase de asignación

3.2.1. Fase de entrenamiento

Para que una imagen pueda ser interpretada de forma fiable es preciso que exista una experiencia previa, que nos permita identificar cada una de las categorías de interés por una serie de rasgos, como son tono, textura, situación o tamaño. Es posible distinguir entre matorral y arbolado, gracias a que éste tiene una textura más rugosa; mientras la separación entre agua y suelos descubiertos se realiza sobre la diferencia de tonalidad. De la misma forma, la clasificación digital se inicia caracterizando los patrones que definen en la imagen las distintas categorías objetivo. Por cuanto se trata de una clasificación basada en los valores numéricos de los píxeles, esta caracterización también debe ser numérica; esto es, se trata de obtener el ND (Nivel Digital), o mejor aún el rango de ND (o valores de reflectancia si la imagen ya está pre-procesada), que identifica a cada categoría, para todas las bandas que intervienen en la clasificación. Por ello, la fase de entrenamiento resulta ciertamente compleja pues se trata de definir con rigor cada una de las categorías que pretendan discriminarse, teniendo en cuenta su propia variabilidad en la zona de estudio.

A continuación, abordaremos los puntos clave de la fase de entrenamiento:

- Selección de las regiones de interés de entrenamiento
- Estudio de la separabilidad de las clases

3.2.1.1. Selección de la región de interés

Como ya hemos indicado anteriormente, necesitamos un conocimiento previo para la realización de este tipo de clasificación. El término región de entrenamiento indica que tales áreas sirven para “entrenar” en el reconocimiento de las distintas categorías. En otras palabras, a partir de ellas el ordenador calcula los niveles digitales que definen cada una de las clases, para luego poder asignar el resto de píxeles de la imagen a una de esas categorías en función de sus niveles digitales. El conjunto de estas zonas de entrenamiento se conoce como región de interés (ROI) [48]. En nuestro caso podemos, con ayuda de una herramienta de nuestro programa, introducir sobre la pantalla los polígonos que definen cada una de las áreas, asignándolas a una de las categorías definidas previamente. Acabada la delimitación de una parcela de entrenamiento, pueden seleccionarse otras áreas para la misma categoría, para así garantizar el poder abarcar mejor la variabilidad de los niveles asociados a cada clase o bien culminar con ella la definición de esa clase.

3.2.1.2. Estudio de la separabilidad de las clases

Finalizada la definición de nuestra región de interés, debemos realizar una serie de cálculos para poder conocer previamente la separabilidad de nuestras clases, esto se conoce con el nombre de cómputo de separabilidad. El primer paso de este cómputo es calcular las estadísticas elementales de cada categoría: media, rango (máximo y mínimo en niveles digitales) y la desviación típica. Realizaremos el cálculo de cada una de estas variables para cada una de las bandas de nuestra imagen original. La segunda parte para realizar nuestro cómputo están relacionadas con la regla de decisión de máxima probabilidad. Así que, la evaluación de la separabilidad de clases ayuda a predecir el resultado de la fiabilidad de la clasificación por máxima probabilidad. Hay dos formas comúnmente empleadas para calcular la separabilidad. Todas estas medidas tienen en cuenta tanto las covarianzas de cada una de las clases en cada una de las bandas, como el vector de medias de cada una de las clases:

- **Divergencia transformada:** da un crecimiento exponencial ponderado con el incremento de las distancias entre clases. Como regla general, si el resultado es mayor que 1.9 se pueden separar las clases. Entre 1.7 y 1.9 la separación es moderadamente buena. Por debajo de 1.7 la separación es mala [49]. Para poder

definir totalmente la divergencia transformada debemos conocer antes la divergencia:

$$D_{ij} = \frac{1}{2} \text{tr} \left((C_i - C_j)(C_i^{-1} - C_j^{-1}) \right) + \frac{1}{2} \text{tr} \left((C_i^{-1} - C_j^{-1})(\mu_i - \mu_j)(\mu_i - \mu_j)^T \right) \quad (3.1)$$

donde:

i, j = son las dos clases comparadas.

C_i = la matriz de covarianza de la clase i.

μ_i = es el vector de las medias de la clase i.

tr = corresponde a la matriz de traza (álgebra matricial).

T = función traspuesta.

Finalmente, la fórmula de la divergencia transformada es la siguiente:

$$TD_{ij} = 200 \left(1 - e^{-\frac{D_{ij}}{8}} \right) \quad (3.2)$$

donde:

i, j = son las dos clases comparadas.

D_{ij} = la divergencia de las clases i e j.

- **Jeffries-Matusita (JM):** Este índice mide la separación entre dos clases en una escala de [0 - 2] [50]. La distancia JM tiene un comportamiento de saturación hacia el valor 2 a medida que aumenta la separación de clases. Para poder definirla matemáticamente debemos conocer primero la distancia de Bhattacharyya, que básicamente, mide la similitud de dos variables de dos distribuciones de probabilidad:

$$B = \frac{1}{8} (\mu_i - \mu_j)^T \left(\frac{C_i + C_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\frac{|C_i + C_j|}{\sqrt{|C_j| |C_i|}} \right) \quad (3.3)$$

donde:

i, j = son las dos clases comparadas.

C_i = la matriz de covarianza de la clase i.

μ_i = es el vector de las medias de la clase i.

$|C_i|$ = es el determinante de la matriz de covarianza.

Ahora, partiendo de la distancia anterior definimos el índice de Jeffries-Matusita

como:

$$JM_{ij} = \sqrt{2(1 - e^{-B})} \quad (3.4)$$

Tanto la divergencia como la distancia Jeffries-Matusita poseen límites inferiores y superiores. Si el valor calculado toma el límite superior se puede decir que esas dos clases son totalmente separables, si el valor es nulo significa que no se pueden separar.

En resumen, lo primero que hemos comentado es cómo hemos realizado el estudio previo para realizar nuestra clasificación supervisada, en base a los resultados obtenidos, y hemos definido una serie de clases y el conjunto de éstas lo hemos llamado región de interés. Para evaluar la separabilidad de nuestras clases hemos realizado una serie de cálculos, una serie de variables estadísticas elementales y unos índices (Jeffries-Matusita y Divergencia transformada).

3.2.2. Fase de asignación

Los principales retos en el procesado de datos de teledetección se orientan a la generación de índices multispectrales o a la aplicación de técnicas de segmentación o clasificación que nos permiten obtener un mapa temático donde se muestra de forma clara el parámetro bajo estudio. El estado del arte en técnicas de clasificación [51], [52], [53], [54] se centra en la aplicación de algoritmos más eficientes como las máquinas de vectores soporte o la clasificación borrosa y otras técnicas novedosas como las basadas en el contexto espacial, textura, información multitemporal, etc. En el ámbito de la alta resolución espacial, también es preciso destacar las novedosas metodologías OBIA (*Object Based Image Analysis*) [55] que interpretan las imágenes no en base a píxeles, sino a los objetos que aparecen en ellas, de forma análoga a como lo hace el Sistema Visual Humano.

En esta fase se trata de asignar cada uno de los píxeles de la imagen a una de las clases previamente seleccionadas. Esta asignación se realiza en función de los ND (Niveles Digitales) de cada píxel, para cada una de las bandas que intervienen en el proceso. Fruto de esta fase será una nueva imagen, cuyos ND expresen la categoría temática a la que se ha adscrito cada uno de los píxeles de la imagen original.

En este apartado se tratan los algoritmos más empleados para realizar este proceso, en el contexto de las aplicaciones de la teledetección. Desde un punto de vista estadístico, todos ellos definen un área de dominio en torno al centro de cada categoría a discriminar, mediante un conjunto de funciones discriminantes. Estas ecuaciones

pueden considerarse, por tanto, como las fronteras que definen cada categoría. El píxel será asignado a una clase, si sus ND se encuentran dentro del área de dominio señalado para dicha clase.

A continuación describiremos en detalle los clasificadores de Máxima Probabilidad (*Maximum Likelihood*) y SVM (*Support Vector Machine*), puesto que serán los empleados en este trabajo.

3.2.2.1. Máxima probabilidad

Este método considera que los niveles digitales en el seno de cada clase se ajustan a una distribución normal. Esto nos permite describir esa categoría por una función de probabilidad a partir de su vector de medias y la matriz varianza-covarianza [56]. El cálculo se realiza para todas las categorías que intervienen en la clasificación, asignando el píxel a aquélla que maximice la función de probabilidad.

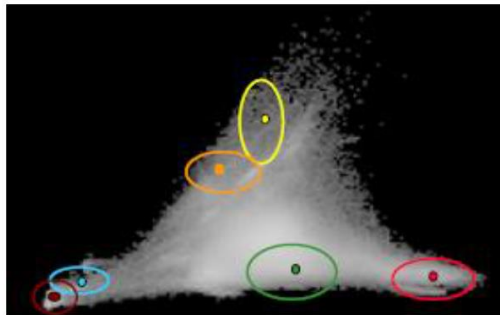


Figura 3.2. Clasificador de máxima probabilidad [57].

El clasificador de máxima probabilidad es el más empleado en teledetección, por su robustez y por ajustarse con más rigor a la disposición original de los datos.

- **Función de discriminación**

La clasificación está basada en el teorema de Bayes [58]:

$$x \in C_i \text{ si } p(C_i|x) > p(C_j|x) \text{ para todo } j \neq i \quad (3.5)$$

donde:

$p(C_i|x)$ = probabilidad condicionada.

Ahora enunciaremos el teorema de Bayes:

$$p(C_i|x) = \frac{p(C_i|x) \cdot p(C_i)}{p(x)} \quad (3.6)$$

donde:

$p(C_i)$ = probabilidad de la clase

$p(x)$ = probabilidad de encontrar algún píxel en la localización x .

Nosotros asumimos que la distribución de probabilidad se corresponde con una distribución normal multivariante:

$$p(x|C_i) = \frac{1}{(2\pi)^{\frac{N}{2}} |\Sigma_i|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(x-m_i)^t \Sigma_i^{-1} (x-m_i)} \quad (3.7)$$

donde:

m_i = vector de medias de la clase C_i

Σ_i = matriz de covarianza de la clase C_i

- **Regla de decisión [59]:**

La regla de decisión divide el conjunto de posibles valores en dos regiones de forma que se minimice la probabilidad de error.

$$x \rightarrow C_i \text{ si } g_i(x) > g_j(x) \text{ para todo } j \neq i \quad (3.8)$$

$$g_i(x) = \ln(p(x|C_i)) + \ln(p(C_i)) \quad (3.9)$$

$$g_i(x) = -\ln(|\Sigma_i|) - (x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (3.10)$$

- **Superficies de decisión**

La superficie de separación espectral entre dos clases i e j viene dada por la expresión:

$g_i(x) = g_j(x) \rightarrow g_i(x) - g_j(x) = 0$, dando lugar a superficies cuadráticas.

$$(x - m_i)^t \Sigma_i^{-1} (x - m_i) \quad (3.11)$$

Los umbrales pueden ser aplicados para asegurar que un píxel tiene un valor mínimo de probabilidad para pertenecer a una cierta clase [60], como se muestra en la Figura 3.3.

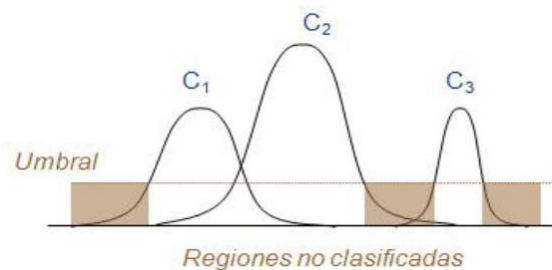


Figura 3.3. Superficies y umbrales de decisión [57].

3.2.2.2. Máquina de Vectores Soporte (SVM)

Las Máquinas de Vectores de Soporte (*Support Vector Machine*) representan a un grupo de algoritmos de aprendizaje automático superiores. El desarrollo de SVM fue provocado inicialmente por la exploración y formalización de aprendizaje del control de la capacidad de la máquina y el sobreajuste de cuestiones [61], [62], [63]. Aunque SVMs se puede decir que ha surgido a finales de 1970 [61], no ha recibido una atención significativa hasta hace pocos años. El atractivo de SVMs es su capacidad para reducir al mínimo el llamado riesgo estructural, o los errores de clasificación, al resolver el problema de clasificación. A diferencia de los métodos convencionales de probabilidad estadística máximos que reducen al mínimo el error de clasificación en un sentido empírico, que está directamente determinado por la distribución de muestras de entrenamiento, el concepto de minimización del riesgo estructural adoptado por SVM es reducir al mínimo la probabilidad de clasificación errónea de un punto de datos inédito extraído al azar de una distribución de probabilidad fija, pero desconocida [62], [63]. Tal propiedad es también diferente de la lógica de decisión límite de formación de las Redes Neuronales Artificiales. Específicamente, SVM siempre encuentra un mínimo global. A lo largo de este apartado veremos como su sencilla interpretación geométrica proporciona un terreno fértil para una mayor investigación [64].

SVMs eran originalmente clasificadores binarios lineales, que asignaban las etiquetas 1 y -1. El funcionamiento básico de SVMs consiste en construir un hiperplano de separación (es decir, un límite de decisión) sobre la base de las propiedades de las muestras de entrenamiento, específicamente su distribución en el espacio de características. Dicho plano de separación está sujeto a la condición de que el margen de separación, entre las muestras de la clase 1 y de la clase -1, se maximiza [61]. En la Figura 3.4, se ilustra esa idea en la cual el hiperplano dibujado con una línea continua separa las dos clases con el máximo margen, mientras que el resto de planos, dibujados con una línea discontinua, no cumplen los requisitos de la minimización del riesgo estructural.

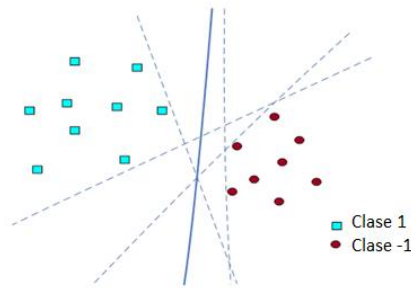


Figura 3.4. Ejemplo de un hiperplano que separa las clases con el máximo margen [61].

Observaremos en las derivaciones matemáticas posteriores, que no necesariamente todas las muestras de entrenamiento contribuyen a la construcción del hiperplano, y que normalmente sólo un subconjunto de muestras de entrenamiento es elegido como vectores de soporte. Este atributo es único para SVMs.

En el campo de la teledetección, la aplicación de SVMs se ha extendido gradualmente y los resultados muestran que generalmente SVM presenta un resultado de clasificación mejorado en comparación con los clasificadores tradicionales como los de máxima verosimilitud [65], [66], [67], [68].

A continuación, mostraremos los conocimientos necesarios que nos permitirán ser capaces de utilizar de forma inteligente SVMs para realizar la clasificación de imágenes de teledetección.

Partiremos de manera introductoria con el caso lineal, seguido por la solución del problema de clasificación no lineal en función de los métodos de la función kernel conocidos. También tendremos en cuenta las formas de lidiar con el caso multiclase y la asignación de parámetros.

- **Clasificación Lineal**

- a) **Caso Separable**

En primer lugar, debemos asumir que hay dos clases linealmente separables. El conjunto de datos de entrenamiento se representa por pares $\{x_i, y_i\}, i = 1, \dots, n, y_i \in \{1, -1\}, x_i \in R^d$, donde x_i son las características multispectrales observadas e y_i es el nivel de información de clase para el caso de entrenamiento i . El nivel es o bien $+1$ o -1 , representando clase 1 y clase 2. El objetivo del clasificador vector soporte es construir un hiperplano óptimo que separe las dos clases de manera que la distancia entre el hiperplano y los puntos de datos de entrenamiento más cercanos de cada una de las clases sea el más grande posible. Esta distancia se denomina margen. El hiperplano puede ser representado por la siguiente función de decisión:

$$w^T x + b = 0 \quad (3.12)$$

donde x es un punto situado en el hiperplano, w es normal al hiperplano, T denota la traspuesta de la matriz, y el parámetro b denota la parcialidad. La distancia perpendicular desde el hiperplano hasta el origen es $|b|/\|w\|$, donde $\|w\|$ es la norma Euclidea de w . En la Figura 3.5, podemos observar el concepto explicado.

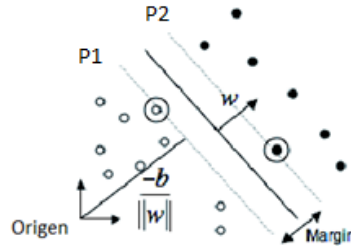


Figura 3.5. Caso linealmente separable [61].

Supongamos que todos los datos de entrenamiento satisfacen las siguientes restricciones:

$$w^T x_i + b \geq +1, \text{ for } y_i = +1$$

$$w^T x_i + b \leq -1, \text{ for } y_i = -1$$

Estas dos ecuaciones pueden combinarse adicionalmente

$$y_i(w^T x_i + b) - 1 \geq 0 \quad (3.13)$$

Se puede entonces implícitamente definir una escala para $(w \times b)$ y generar dos hiperplanos canónicos (P1 y P2 en la Figura 3.5), concretamente, $w^T x_i + b = 1$ para los puntos más cercanos de entrenamiento en un lado del hiperplano con distancia normal w y perpendicular desde el origen $|1 - b|/\|w\|$, y en segundo lugar, $w^T x_i + b = -1$ para los puntos más cercanos de entrenamiento por el otro lado con distancia perpendicular desde el origen $|-1 - b|/\|w\|$. Estos puntos de entrenamiento se denominan vectores de soporte. En otras palabras, esos vectores son fundamentales para el establecimiento del hiperplano óptimo de separación. En consecuencia, el margen entre esos dos hiperplanos es $2/\|w\|$. La maximización de este margen a su vez conduce a:

$$\min \left\{ \frac{\|w\|^2}{2} \right\} \quad (3.14)$$

sujeto a las restricciones de desigualdad de la ecuación combinada descrita anteriormente. Dicha ecuación se hace más fácil de manejar si se utiliza una formulación de Lagrange:

$$L_{\text{primal}} = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i (y_i (w \times x_i + b) - 1) + \sum_{i=1}^n \alpha_i \quad (3.15)$$

donde α_i son multiplicadores positivos de Lagrange. Debemos minimizar L_{primal} con respecto a w y b . A partir del teorema de Wolfe [68], se puede tomar la derivada de L_{primal} con respecto a b y w para obtener

$$\frac{\partial L_{\text{primal}}}{\partial b} = 0 \rightarrow \sum_{i=1}^n \alpha_i y_i = 0 \quad (3.16)$$

$$\frac{\partial L_{\text{primal}}}{\partial w} = 0 \rightarrow w = \sum_{i=1}^n \alpha_i y_i x_i \quad (3.17)$$

y sustituyendo en la formulación primal para obtener el dual Wolfe de Lagrange:

$$L_{\text{dual}} = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \alpha_i \alpha_j y_i y_j x_i \times x_j \quad (3.18)$$

El entrenamiento de SVM ahora incluye la maximización de esta nueva ecuación con respecto a $\alpha_i \geq 0$, sujeto a la restricción mostrada en la ecuación de la derivada parcial respecto a b , con la solución dada por la ecuación de la derivada parcial con respecto a w . Advertir de que hay un multiplicador de Lagrange α_i para cada punto de entrenamiento. Estos puntos con $\alpha_i > 0$ son vectores soporte, y se encuentran en uno de los hiperplanos paralelos P1 o P2 (Figura 3.6). Todos los demás puntos de entrenamiento tienen $\alpha_i = 0$ y se encuentran ya sea en P1 o P2. Específicamente, las soluciones para w y b se formulan como

$$w = \sum_{i=1}^{\text{nsv}} \alpha_i y_i x_i \quad (3.19)$$

$$b = -\frac{1}{2} w \times (x_r + x_s) \quad (3.20)$$

donde nsv denota el número de vectores soporte, x_r es el vector soporte perteneciente a la clase $y_r = 1$, y $y_s = -1$ para x_s . Debemos tener en cuenta que la parcialidad b puede ser calculada utilizando todos los vectores de soporte en el margen de preocupación de estabilidad.

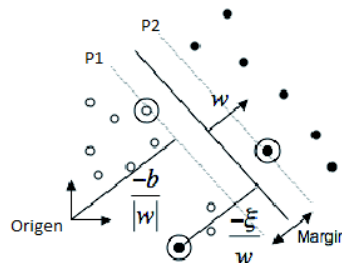


Figura 3.6. Hiperplanos lineales para el caso parcialmente separable [61].

En consecuencia, en el caso de un problema de clasificación de dos clases, la regla de decisión mostrada en la ecuación 3.16, que separa las dos clases de información, se puede derivar como:

$$f(x) = \text{sign}(\sum_{i=1}^{n_{sv}} \alpha_i y_i (x \times x_i) + b) \quad (3.21)$$

Esta formulación del problema de optimización de SVM se denomina formulación de margen duro, ya que no se permiten errores de entrenamiento. Todas las muestras de entrenamiento satisfacen la desigualdad $y_i \times f(x_i) \geq 1$. En ciertos casos, se puede requerir el hiperplano de separación al pasar por el origen por elegir un $b=0$ fijo. Esta variante se denomina el margen duro de SVM sin umbral. En este caso, el problema de optimización sigue siendo el mismo que el de la ecuación 3.21 excepto que la restricción $\alpha_i \times y_i = 0$ de la ecuación 3.20 ya no existe.

b) Caso No Separable

Clases de información derivadas de los datos de teledetección no suelen ser totalmente separadas por límites lineales. Por tanto, las limitaciones de la ecuación 3.13 no pueden ser satisfechas en la práctica, por lo que las variables de holgura o laxitud ξ_i , $i=1, \dots, n$, que son proporcionales a alguna medida de costo, se introducen para relajar las restricciones [69]. Esto a veces se llama método del margen suave. Podemos observar esta idea en la Figura 3.6. Cuando se incluyen las variables de holgura, la ecuación 3.13 se convierte en

$$y_i(w^T \times x_i + b) \geq 1 - \xi_i; \xi_i \geq 0, \forall i. \quad (3.22)$$

El problema óptimo se resuelve entonces mediante:

$$\min \left\{ \frac{\|w\|^2}{2} + C \sum_{i=1}^n \xi_i \right\} \quad (3.23)$$

donde C es el parámetro de penalización o regularización que debe ser determinado por el usuario. Un mayor C significa “asignar una penalización superior a los errores”. La primera parte de la ecuación 3.23 tiene como objetivo maximizar el margen, mientras que la segunda parte busca penalizar a las muestras de entrenamiento ubicadas en el lado equivocado de la frontera de decisión. En otras palabras, un mayor valor de C se corresponde con un sobreajuste a los datos de entrenamiento y por lo tanto reduce la capacidad de generalización. Belousov [70], sin embargo, muestra que SVMs puede mostrar un alto grado de robustez frente a variaciones en los valores de los parámetros.

La formulación primal de Lagrange de la ecuación 3.11 se convierte entonces

$$L_{\text{primal}} = \frac{1}{2} \|w\|^2 - \sum_{i=1}^n \alpha_i (y_i (w^T \times x_i + b) - 1 + \xi_i) + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \mu_i \xi_i \quad (3.24)$$

donde μ_i son los multiplicadores de Lagrange introducidos para hacer cumplir la positividad de la variable de holgura ξ_i . La solución de la ecuación 3.24 viene

determinada por los saddles points de Lagrange (puntos en los que el potencial se curva hacia arriba en una dirección y hacia abajo en la otra), minimizando con respecto a w , ξ y b , y maximizando con respecto a $C \geq \alpha_i \geq 0$ y $\mu_i \geq 0$. Los multiplicadores de Lagrange α_i ahora tienen un límite superior de C . El límite superior $\alpha_i \leq C$ corresponde a un límite superior en la fuerza que se permite cualquier vector de apoyo dado a ejercer sobre los hiperplanos $P1$ y $P2$. Por último, en el caso no separable la formulación del dual Wolfe de Lagrange sigue siendo el mismo que el mostrado en la ecuación 3.18.

- **Clasificación No Lineal y funciones Kernel**

En muchos casos en los que un hiperplano lineal es incapaz de separar las clases apropiadamente, SVMs adapta una estrategia bastante antigua [71], que asigna los datos de entrada sin procesar en un espacio dimensional superior a fin de mejorar la separabilidad entre clases. Sólo entonces, SVMs lleva a cabo la manipulación dentro de este espacio de características recién asignado. Este método se denota como SVMs no lineal.

- a) **SVM No Lineal**

Para obtener una superficie de decisión no lineal, Boser [72] propone que un vector de características, $x_i \in \mathbb{R}^d$, esté asignado en un espacio dimensional superior euclídeo (o una generalización del espacio euclídeo llamado espacio de Hilbert [73], [74]), con el fin de difundir la distribución de las muestras de entrenamiento de una manera que facilita la instalación de un hiperplano lineal. Específicamente, las muestras de entrenamiento se proyectan en un mayor espacio dimensional (\mathcal{H}), a través de una función vector no lineal de mapeo llamado $\phi: \mathbb{R}^d \rightarrow \mathcal{H}$. Recordemos la ecuación 3.21, que puede ser derivada adicionalmente como

$$f(x) = \text{sign}(\sum_{i=1}^{n_{sv}} \alpha_i y_i \phi(x) \cdot \phi(x_i) + b) \quad (3.25)$$

La carga computacional de $(\phi(x) \cdot \phi(x_i))$ puede ser bastante elevada en un espacio dimensional superior. Vapnik [61] propone una alternativa para reducir la carga computacional a través de una función kernel definida positiva indicada como $K(x, y)$, tal que $K(x, y) = \phi(x) \cdot \phi(y)$. Para ilustrar aún más esta relación, permitimos $K(x, y) = (x \times y)^2$, donde $x = (x_1, x_2)$, e $y = (y_1, y_2)$. Se puede entonces extender la función de kernel como

$$\begin{aligned} K(x, y) &= (x_1 y_1 + x_2 y_2)^2 \\ &= x_1^2 y_1^2 + 2x_1 y_1 x_2 y_2 + x_2^2 y_2^2 \end{aligned}$$

$$= (x_1^2, x_2^2, \sqrt{2}x_1x_2) \times (y_1^2, y_2^2, \sqrt{2}y_1y_2) = \phi(x) \cdot \phi(y) \quad (3.26)$$

Claramente, acorde a la ecuación 3.26, mediante el uso de la función kernel $K(x, y) = (x \times y)^2$ en lugar de $\phi(x) \cdot \phi(y) = (x_1^2, x_2^2, \sqrt{2}x_1x_2) \times (y_1^2, y_2^2, \sqrt{2}y_1y_2)$, la representación del dato es enormemente simplificada y la carga computacional se reduce. Se necesitaría solamente utilizar la función K en el algoritmo de entrenamiento, y probablemente no hay necesidad de conocer explícitamente lo que ϕ es en realidad. Más adelante, veremos en detalle otras funciones de kernel. El problema de optimización de la ecuación 3.18 puede entonces ser reescrito como:

$$L_{\text{dual}} = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i \times x_j) \quad (3.27)$$

y la regla de decisión expresada en la ecuación 3.11 ahora se generaliza en:

$$f(x) = \text{sign}(\sum_{i=1}^{nsv} \alpha_i y_i K(x, x_i) + b) \quad (3.28)$$

De hecho, el papel de la función kernel consiste en el mapeo de las muestras de entrenamiento en un espacio de dimensiones superiores donde las muestras de entrenamiento pueden extenderse de forma más separada. En la Figura 3.7 vemos ilustrada esta idea en la cual, originalmente, los datos eran linealmente no separables en un espacio de dos dimensiones, pero una vez mapeados en un espacio de características de dimensiones superiores $\phi()$, los datos pueden ser más fácilmente separables en términos de hiperplano.

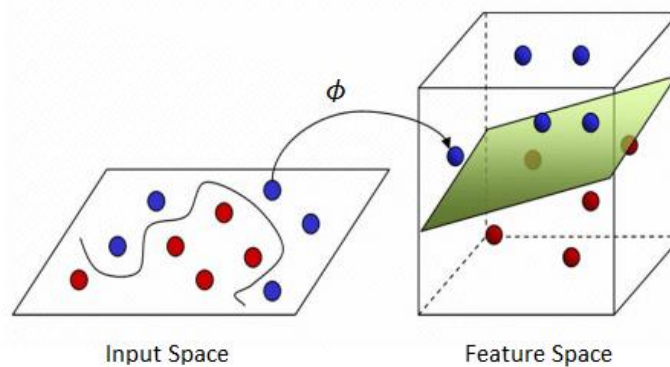


Figura 3.7. Caso linealmente no separable, en el cual SVM mapea los datos en bruto en una dimensión superior para aumentar la separabilidad entre clases [75].

Debemos tener en cuenta que la función kernel se utiliza para adaptar SVMs que debe satisfacer la condición de Mercer, es decir, la función kernel debe satisfacer la restricción

$$\int \int K(x, y) g(x) g(y) d(x) d(y) \geq 0 \quad (3.29)$$

donde $g(x)$ es una función cuadrada integrable así como garantizar que el par de mapeo $\{\phi, \mathcal{H}\}$ existe, y consecuentemente el proceso de entrenamiento converge [26]. Debemos tener en cuenta también que para el caso no separable, los SVMs de márgenes suaves con penalización cuadrática de errores pueden considerarse como un caso especial de la versión de margen duro con el kernel modificado $K = K + I/C$, donde I representa a la matriz identidad [69].

b) Funciones Kernel

Anteriormente hemos analizado que la función kernel tiene como propósito permitir operaciones que se realizarán en el dominio actual (por ejemplo, $K(x, y)$ en la ecuación 3.28) en lugar de en un espacio dimensional de características de alto nivel (por ejemplo, $\phi(x) \times \phi(y)$ en la ecuación 3.25). Esto proporciona los medios inteligentes para resolver el problema computacional causado por la alta dimensionalidad. Vamos a introducir cinco opciones comunes de funciones de kernel:

Polinomial (homogéneo):

$$K(x_i, x_j) = \gamma(x_i \times x_j)^d \quad (3.30)$$

Polinomial (inhomogéneo):

$$K(x_i, x_j) = (\gamma(x_i \times x_j) + \delta)^d, \gamma > 0, \delta > 0 \quad (3.31)$$

Función de base radial:

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (3.32)$$

Función Gaussiana de base radial:

$$K(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2), \gamma > 0 \quad (3.33)$$

Sigmoideo:

$$K(x_i, x_j) = \tanh(\gamma(x_i \times x_j) - \delta), \gamma > 0, \delta > 0 \quad (3.34)$$

Huang y otros autores [65] muestran que ambas ecuaciones polinomiales son métodos populares para el mapeo no lineal con orden polinomial $d > 1$, y hay una obvia tendencia a mejorar la precisión conforme d se incrementa. Esto también sugiere que el valor de d tiene diferente impacto en el rendimiento del kernel cuando son usados diferentes números de variables de entrada. Cuando se trata del kernel de función de base radial, Schölkopf muestra que la función Gaussiana de base radial obtiene excelentes resultados comparados con la función clásica de base radial. La función

kernel sigmoidea, puede ser considerada como un tipo particular de la red neuronal sigmoidea. Debemos darnos cuenta de que la función kernel sigmoidea puede no satisfacer la condición de Mercer para ciertos valores de γ y δ [62].

Con estas descripciones está claro que el rendimiento de SVMs es completamente relativo a la elección de la función kernel. Aunque algunos estudios han trabajado en la elección de kernels adecuados basados en el conocimiento previo [64], la mejor elección del kernel para un problema dado todavía requiere una mayor investigación. Una vez se determina el tipo de kernel, el usuario tiene que elegir cuidadosamente los parámetros del kernel correspondiente (parámetros como d y γ) y el error de penalización para obtener unos resultados en la clasificación lo más satisfactorios posibles. Es muy difícil determinar cuál es el patrón de decisión óptimo. Claramente dichos valores dependerán del caso de estudio con la clara evidencia de que la elección de los parámetros afectará a los resultados de la clasificación.

En este proyecto, nos centraremos en el uso del kernel polinomial y la función de kernel de base radial ya que, según las investigaciones realizadas, son las dos funciones kernel que mejores resultados ofrecen en teledetección [76]. De hecho, estudiaremos con mayor énfasis la función de base radial ya que diversos autores han concluido que ofrece resultados superiores a los del kernel polinomial.

- **Determinación de Parámetros**

La principal limitación en el uso de SVMs es la necesidad de seleccionar valores de parámetros adecuados a fin de mejorar la capacidad de las SVMs bien entrenadas para predecir con precisión datos desconocidos. Para lograr el objetivo de seleccionar unos valores apropiados para los parámetros, primero debemos garantizar que la estimación de errores de validación es imparcial. Entonces, usando esa estimación del error imparcial, uno puede realizar ajustes de parámetros que mejoran la capacidad de generalización del clasificador. En esta sección se ilustran formas de determinar esa estimación del error de manera imparcial, y esto es seguido por una consideración de dos tipos de metodología de ajuste de parámetros, conocidos como la red de búsqueda (*grid search*) y la técnica de descenso de gradiente, respectivamente.

- a) ***T-Fold Cross-Validation***

Las muestras terrestres disponibles se asignan al azar a uno de dos conjuntos, un conjunto que se utiliza para fines de entrenamiento y el otro tratado como datos de

prueba desconocidos para fines de validación. Si las muestras disponibles son suficientes, la estimación del error de validación será menos sesgada. En el caso de SVMs, la estimación del error se puede escribir como:

$$P_{\text{error}}^t = \frac{1}{t} \sum_{i=1}^t \Gamma(-y_i f(x_i)) \quad (3.35)$$

donde t es el número de muestras para validación, $f(x_i)$ es la función de decisión como la tratada en la ecuación 3.28, y Γ es la función escalón, $\Gamma(\beta) = 1$ cuando $\beta > 0$ y $\Gamma(\beta) = 0$ en cualquier otro caso. Los parámetros que permitan alcanzar la máxima precisión de validación serían los candidatos ideales para especificar el clasificador final. Dado que los límites de decisión SVMs están definidos por los vectores de soporte, el entrenamiento y validación de una sola prueba no pueden reflejar con realismo la distribución real de las muestras no clasificadas, especialmente en los casos en que el número de muestras es bastante limitado. En otras palabras, los resultados de SMVs que permitan la más alta precisión de entrenamiento podrían no generalizarse con éxito para las muestras desconocidas. La técnica llamada *t-fold cross-validation* es generalmente implementada para solucionar este problema.

En el enfoque de *t-fold cross-validation*, las muestras de entrenamiento son separadas a partes iguales en un número de subconjuntos t definido por el usuario. Durante la fase de entrenamiento, el usuario escoge uno de los t subconjuntos para la validación y los restantes $(t-1)$ subconjuntos para el entrenamiento. Este proceso se repite secuencialmente hasta que todos los conjuntos de datos hayan sido probados. Los errores de clasificación calculados a partir de conjuntos de validación son entonces promediados durante las t pruebas. Este método es también denominado como *leave-one-out cross-validation*. La precisión de validación cruzada es el porcentaje de las muestras de validación que han sido correctamente clasificadas, y esta medida proporciona una estimación más objetiva y precisa de la precisión de clasificación que el método *one-trial*. Según el estudio realizado por Luntz y Brailovsky [77], el procedimiento *leave-one-out* puede dar una estimación casi imparcial de la generalización de error esperado con respecto a las muestras de validación.

b) Bound on Leave-One-Out Error

Aunque el uso de *t-fold leave-one-out* de validación cruzada para estimar el error generalizado es generalmente recomendado, puede llegar sin embargo a ser muy costoso para calcular realmente, ya que requiere la ejecución de los t tiempos de los algoritmos

de entrenamiento. La estrategia alternativa es, por lo tanto, la de límite superior (*upper bound*) o aproximar el error por una estimación que es más fácil de calcular. Si f^{all} denota el entrenamiento de SVM para formar la frontera de decisión utilizando todas las muestras de entrenamiento, y f^i la frontera de decisión obtenida cuando la muestra i es eliminada, se puede derivar la generalización de error esperada para la SVM como:

$$E[P_{\text{error}}^{t-1}] = \frac{1}{t} \sum_{i=1}^t \Gamma(-y_i f^i(x_i)) \quad (3.36)$$

donde P_{error}^{t-1} indica el actual error para una SVM entrenado en $t-1$ muestras, $E[P_{\text{error}}^{t-1}]$ es la expectativa del error real sobre todas las opciones de conjunto de entrenamiento de tamaño $t-1$, y la expectativa E se toma sobre una elección aleatoria de muestras de entrenamiento. El lado derecho de la ecuación 3.20 se puede derivar adicionalmente como:

$$\frac{1}{t} \sum_{i=1}^t \Gamma(-y_i f^i(x_i)) = \frac{1}{t} \sum_{i=1}^t \Gamma(-y_i f^{\text{all}}(x_i) + y_i (f^{\text{all}}(x_i) - f^i(x_i))). \quad (3.37)$$

Asumir que el término $y_i (f^{\text{all}}(x_i) - f^i(x_i))$ está acotado por U_i . Se puede entonces expresar el siguiente límite superior en el error *leave-one-out* como:

$$E[P_{\text{error}}^{t-1}] \leq \frac{1}{t} \sum_{i=1}^t \Gamma(U_i - 1). \quad (3.38)$$

Varios estudios desarrollan estimaciones de la cota de error de diferentes maneras como se aborda a continuación.

Basándose en la teoría de SVMs, se sabe que la eliminación de un no vector de soporte del conjunto de entrenamiento no cambia la solución calculada por el SVM (es decir, $U_i = 0$, para el no vector de soporte x_i), en consecuencia Vapnik [62] deriva la cota del error *leave-one-out* de SVMs como:

$$E[P_{\text{error}}^{t-1}] \leq \frac{E[\text{número de vectores soporte}]}{t}. \quad (3.39)$$

Para SVMs sin umbral b (ecuación 3.7), Jaakkola y Haussler [78] propusieron otra cota de error *leave-one-out*:

$$E[P_{\text{error}}^{t-1}] \leq \frac{1}{t} \sum_{i=1}^t \Gamma(\alpha_i K(x_i, x_j) - 1) \quad (3.40)$$

donde α_i es el multiplicador de Lagrange para el vector soporte i y $K(x_i, x_j)$ es la función kernel. Algunos autores [79] adoptaron el concepto de la teoría de respuesta lineal para proponer el siguiente límite de error de generalización bajo el supuesto de que el conjunto de vectores de soporte no cambia al retirar la muestra i :

$$E[P_{\text{error}}^{t-1}] \leq \frac{1}{t} \sum_{i=1}^t \Gamma\left(\frac{\alpha_i}{K_{\text{SV}}^{-1}} - 1\right) \quad (3.41)$$

donde K_{SV}^{-1} es la matriz de productos escalares entre los vectores soporte.

Otra expresión que define el límite del error *leave-one-out* está basada en el concepto de esfera y el radio. Sea $\phi(x_1), \phi(x_2), \dots, \phi(x_t)$ los mapeos de las asignaciones de las muestras de entrenamiento en el espacio de características (tener en cuenta que hay un total de t conjuntos, y donde x_i representa el conjunto de entrenamiento con el conjunto i th excluido), y acostado dentro de una esfera de radio R . Aquí, m es el margen entre dos hiperplanos. Vapnik [63] sugirió el siguiente límite para la generalización del procedimiento del error *leave-one-out*:

$$E[P_{\text{error}}^{t-1}] \leq \frac{1}{t} \frac{R^2}{m^2} \quad (3.42)$$

donde el radio R^2 puede ser derivado en términos de maximización de la siguiente *Wolfe dual* [64]:

$$R^2 = \sum_{i=1}^t \beta_i K(x_i \times x_i) - \sum_{i=1}^t \sum_{j=1}^t \beta_i \beta_j K(x_i \times x_j) \quad (3.43)$$

sujeito a:

$$\sum_{i=1}^t \beta_i = 1, \quad (3.44)$$

$\beta_i \geq 0$ para todo i . Tener en cuenta que la función de kernel $K(x_i \times x_j)$ se utiliza de nuevo para reemplazar $\phi(x_i) \times \phi(x_j)$.

Bajo la suposición de que el conjunto de vectores soporte sigue siendo el mismo durante el procedimiento *leave-one.out*, Vapnik y Chapelle [80] deducen la siguiente estimación de error basada en el concepto de la duración de vectores soporte:

$$E[P_{\text{error}}^{t-1}] \leq \frac{1}{t} \sum_{i=1}^t \Gamma(\alpha_i S_i^2 - 1) \quad (3.45)$$

donde S_i es la distancia Euclidea entre la muestra $\phi(x_j)$ y el conjunto Ω_i , donde Ω_i es definido como:

$$\Omega_i = \left\{ \sum_{i \neq j, \alpha_j > 0} \lambda_j \phi(x_j) \right\} \quad (3.46)$$

sujeito a:

$$\sum_{i \neq j} \lambda_j = 1 \text{ y } \forall i - j, \alpha_j + y_i y_j \alpha_i \lambda_j \geq 0. \quad (3.47)$$

Observar que los valores de λ_i menores que 0 están permitidos. El usuario puede entonces intentar minimizar el límite del error mediante el uso de las fórmulas del límite del error.

c) Grid Search

La cuestión, llegados a este punto, es cómo se busca la manera óptima de elegir los valores adecuados de los parámetros. Los parámetros a determinar, como ya hemos visto anteriormente, son los parámetros relacionados con la función kernel. Un método llamado red de búsqueda o búsqueda mallada (*grid search*) se puede implementar para estimar los valores de los parámetros [81]. El concepto de la búsqueda en red es similar a la búsqueda exhaustiva en un espacio de soluciones. El usuario comienza con una elección aleatoria de valores de parámetros que son las entradas del clasificador para evaluar el rendimiento de *t-fold cross-validation*. Luego, se aumentan o disminuyen los valores de los parámetros, y el rendimiento es reevaluado hasta que todos los parámetros escogidos hayan sido evaluados. Si empleásemos por ejemplo la función kernel de base radial, hay que determinar los posibles pares de valores de los parámetros (γ, C) y comprobar el rendimiento correspondiente. El par de parámetros (γ, C) que produzca la mejor precisión de validación cruzada es entonces seleccionado. Tradicionalmente, una red de búsqueda sobre un espacio acotado (x, y) comienza en una esquina de la red y se evalúa en cada punto de la cuadrícula separados por un valor δ hasta que se alcanza la esquina opuesta. Los valores de parámetros que generen la mayor precisión de clasificación son entonces seleccionados. Para llevar a cabo una completa red de búsqueda, usar un método de paso a paso puede no ser práctico ya que las exigencias computacionales son considerables. Una alternativa es evaluar inicialmente los valores de los parámetros a través de una malla gruesa (mayor δ) y luego centrar la búsqueda con una malla fina (menor δ) que consiste en una pequeña parte de la malla gruesa.

d) Gradient Descent Method

Si el número de parámetros excede de dos, el método *grid search* se convierte en inviable. Se reconoce que como el número de parámetros crece secuencialmente, el espacio de soluciones se expande exponencialmente. En el siguiente párrafo introduciremos la técnica del descenso de gradiente para hacer frente a casos en los que son introducidos un gran número de parámetros [82]. Una vez que el tema de la selección de parámetros se resuelve, es posible que otros kernels posiblemente más complejos puedan ser introducidos a efectos de clasificación.

El enfoque de descenso de gradiente busca en el espacio de parámetros comenzando en un punto aleatorio definido por valores aleatorios de los parámetros y procede cambiando los valores de los parámetros de acuerdo con la pendiente (derivada) y la dirección de la pendiente de la superficie de error. La posición en la superficie de error se define por valores de parámetros, y el valor de la superficie de error en una posición dada es el error de clasificación obtenido con esos parámetros.

- **Clasificación Multiclase**

SVMs fueron originalmente diseñados para clasificaciones binarias, es decir, un SVM sólo podía separar dos clases, sin embargo la mayoría de las aplicaciones de teledetección requieren tratar con múltiples clases. Diversos enfoques, incluido *one-against-one*, *one-against-others*, estrategias de gráficos dirigidos acíclicos (DAG), y SVM multiclases han sido propuestos para la aplicación de SVMs a las clasificaciones multiclase.

- a) ***One-Against-One, One-Against-Others, and DAG***

La estrategia de clasificación uno contra uno (*one-against-one*) fue propuesta por Knerr [83] quien utilizó comparaciones por pares entre todas las k clases. Por lo tanto, cada clasificador está entrenado exclusivamente en dos de las k clases. Todas las posibles combinaciones de clases uno contra uno se evalúan del conjunto de entrenamiento de k clases. Se necesitará un total de $k(k-1)/2$ clasificadores para realizar la tarea. La regla de decisión para el etiquetado de un píxel es en términos de la votación por mayoría. Al píxel se le dará la etiqueta de la clase que tiene el mayor número de votos. En caso de empate, una estrategia de desempate puede ser adoptada. Uno puede seleccionar al azar una de las clases que están vinculadas, o referirse a la clase del píxel vecino más cercano para romper el empate.

La estrategia uno contra otros (*one-against-others*) es la construcción de un total de k SVMs para las k clases. Cada uno de los SVMs es entrenado para clasificar una clase contra todas las demás clases. Al píxel X entonces, se le dará la etiqueta de la clase con el mayor valor de $w_i\phi(x) + b_i, I = 1$ a k [62].

Platt [84] propone el uso de SVMs de grafo dirigido acíclico (DAG) para la clasificación multiclase. La fase de entrenamiento de DAG-SVM es la misma que en la del método *one-against-one*, es decir, construyendo $k(k-1)/2$ clasificadores, donde k es el número de clases. Esos $k(k-1)/2$ clasificadores forman entonces los nodos internos del

grafo con cada nodo siendo un SVM binario usando las clases i th y j th. Empezando por la raíz del DAG, una muestra de ensayo clasificada es comparada en el nodo y entonces, dependiendo del valor de salida del nodo, dirigida al nodo izquierdo o derecho en la siguiente capa, hasta que se alcanza una hoja. La elección del orden de clase para formar un DAG-SVM puede ser arbitraria sin afectar significativamente la precisión de clasificación. Una ventaja en el uso de DAG-SVMs es la mejora sustancial tanto en el tiempo de entrenamiento como en el de evaluación en comparación con los algoritmos *one-against-one* y *one-against-others* [84]. La Figura 3.8 muestra cómo sería un DAG-SVM para 4 clases (A, B, C y D), en total han sido construidos 6 clasificadores binarios. Se puede observar que en cada nivel, el DAG-SVM expulsa una de las clases y luego se dirige hacia el nodo que contiene el par de clases supervivientes hasta que se logre una decisión final.

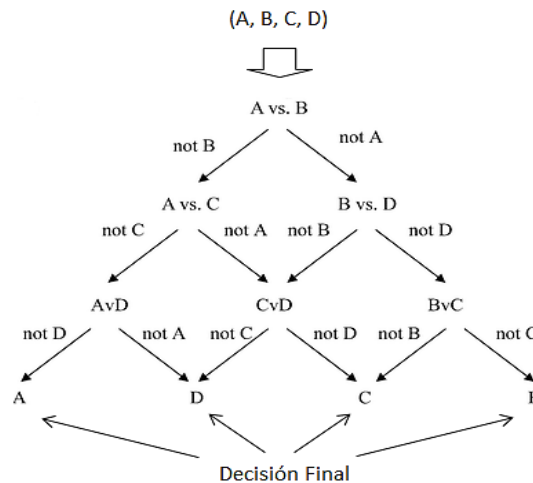


Figura 3.8. Ejemplo de DAG-SVM [84].

b) SVMs Multiclase

i. Enfoque de Vapnik

Vapnik [63] ha desarrollado una aproximación para los problemas de clasificación SVM multiclase. Dicho SVM multiclase es similar al enfoque *one-against-others* mediante el uso de un único enfoque de optimización. Vamos a suponer que tenemos un total de k clases. El SVM multiclase construye k reglas de dos clases en las cuales el hiperplano i th $w_i^T \phi(x) + b$ separa las muestras de entrenamiento de la clase i , del resto de las clases. Por lo tanto, aunque hay k funciones de decisión o hiperplanos, la solución puede ser obtenida mediante la resolución de un problema. La formulación de SVM multiclase se expresa como:

$$\min \left\{ \frac{1}{2} \sum_{i=1}^k w_i^T w_i + \sum_{j=1}^n \sum_{i \neq y_j} \xi_j^i \right\} \quad (3.48)$$

bajo las restricciones:

$$\begin{aligned} w_{y_j}^T \phi(x_j) + b_{y_j} &\geq w_i^T \phi(x_j) + b_i + 2 - \xi_j^i \\ \xi_j^i &\geq 0, j = 1, \dots, n; i \in \frac{\{1, \dots, k\}}{y_j} \end{aligned} \quad (3.49)$$

donde n es el número de muestras de entrenamiento. La función de decisión es entonces

$$\arg \max_{i=1 \text{ to } k} (w_i^T \phi(x) + b_i) \quad (3.50)$$

Hsu y Lin [85] muestran que la derivación de la solución a la ecuación 3.42 requiere que se seleccionen más de k variables en cada iteración, lo cual no es una condición fácil de cumplir. Por lo tanto, para asegurar que una solución válida se obtiene, el elemento $\sum_{i=1}^k b_i^2$ es introducido en la ecuación 3.48 como:

$$\begin{aligned} \min \left\{ \frac{1}{2} \sum_{i=1}^k (w_i^T w_i + b_i^2) + C \sum_{j=1}^n \sum_{i \neq y_j} \xi_j^i \right\} = \\ \min \left\{ \frac{1}{2} \sum_{i=1}^k [w_i^T \quad b_i] \times \begin{bmatrix} w_i \\ b_i \end{bmatrix} + C \sum_{j=1}^n \sum_{i \neq y_j} \xi_j^i \right\} \end{aligned} \quad (3.51)$$

bajo las restricciones:

$$\begin{aligned} [w_{y_j}^T \quad b_{y_j}] \times \begin{bmatrix} \phi(x_j) \\ 1 \end{bmatrix} &\geq [w_i^T \quad b_i] \times \begin{bmatrix} \phi(x_j) \\ 1 \end{bmatrix} + 2 - \xi_j^i \\ \xi_j^i &\geq 0, j = 1, \dots, n; i \in \frac{\{1, \dots, k\}}{y_j}. \end{aligned} \quad (3.52)$$

La ecuación 3.51 puede transferirse a la siguiente formulación [85]:

$$\begin{aligned} \min \left\{ \sum_{r=1}^n \sum_{s=1}^n \left(\frac{1}{2} c_s^{y_r} A_r A_s - \sum_{i=1}^k \alpha_r^i \alpha_s^{y_r} + \frac{1}{2} \sum_{i=1}^k \alpha_r^i \alpha_s^i \right) (K_{r,s} + 1) - \right. \\ \left. 2 \sum_{r=1}^n \sum_{i=1}^k \alpha_r^i \right\} \end{aligned} \quad (3.53)$$

con $0 \leq \alpha_r^i \leq C$, $\alpha_r^{y_r} = 0$, donde:

$$A_r = \sum_{i=1}^k \alpha_r^i, \quad c_s^{y_r} = \begin{cases} 1 & \text{si } y_r = y_s \\ 0 & \text{si } y_r \neq y_s \end{cases}. \quad (3.54)$$

La función de decisión se convierte entonces en:

$$f(x) = \arg \max_{i=1, \dots, k} \left(\sum_{r=1}^n (c_r^i A_r - \alpha_r^i) \times (K(x_r, x) + 1) \right). \quad (3.55)$$

ii. Metodología de Crammer y Singer

Crammer y Singer [86] proponen un enfoque para la solución de los problemas de clasificación multiclase de SVMs en base a la siguiente función de optimización:

$$\min \left\{ \frac{1}{2} \sum_{i=1}^k x_i^T x_i + C \sum_{j=1}^n \xi_j \right\} \quad (3.56)$$

bajo la restricción de que:

$$w_{y_j}^T \phi(x_j) - w_i^T \phi(x_j) \geq e_j^i - \xi_j \quad j = 1, \dots, n \quad (3.57)$$

donde $e_j^i = 1 - \delta_{y_j, i}$, $\xi_j \geq 0$, y $\delta_{y_j, i}$ es definida como:

$$\delta_{y_j, i} = \begin{cases} 1 & \text{si } y_j = i \\ 0 & \text{si } y_j \neq i \end{cases} \quad (3.58)$$

La función de decisión es:

$$\arg \max_{i=1, \dots, k} = w_i^T \phi(x) \quad (3.59)$$

La formulación dual de la ecuación 3.56 es:

$$\min f(\alpha) = \frac{1}{2} \sum_{r=1}^n \sum_{s=1}^n K_{r,s} \bar{\alpha}_r^T \bar{\alpha}_s + \sum_{r=1}^n \bar{\alpha}_r^T \bar{e}_i \quad (3.60)$$

bajo la restricción de que:

$$\sum_{i=1}^k \alpha_r^i = 0, \quad r = 1, \dots, n, \quad \alpha_r^i \leq C, \text{ si } y_i \neq m, \text{ y } \alpha_r^i \leq 0, \text{ por otra parte}$$

donde:

$$\bar{\alpha}_r = [\alpha_r^1, \dots, \alpha_r^k]^T, \text{ y } \bar{e}_r = [e_r^1, \dots, e_r^k]^T. \quad (3.61)$$

Entonces se obtiene:

$$w_i = \sum_{r=1}^n \alpha_r^i \phi(x_r) \quad (3.62)$$

y la función de decisión en la ecuación 3.59 puede ser escrita como:

$$\arg \max_{i=1, \dots, k} = \sum_{r=1}^n \alpha_r^i K(x_r, x). \quad (3.63)$$

Para concluir este punto, disponemos de los conocimientos suficientes para determinar que la metodología de Crammer y Singer es más simple que la de Vapnik. Con la comparativa ofrecida por Hsu y Lin, ambos SVMs multiclase sufren el inconveniente de la convergencia lenta y la precisión de clasificación que ofrecen es menos ideal que la generada por los enfoques one-against-one y DAG.

3.2.3. Fase de Post-procesado

Tras la fase de asignación, a la entrada de nuestra etapa de procesado tenemos una imagen categorizada. En esta etapa, aplicaremos a nuestra imagen una serie de filtrados para intentar mejorar los resultados finales. Existen múltiples filtros utilizados pero nosotros nos centraremos en el *Majority*. La técnica *Majority* de postclasificación emplea un filtro en forma de ventana que se mueve recorriendo toda la imagen, y asigna al píxel del centro de la ventana la clase de la mayoría de sus vecinos dentro de ésta [46]. Este suavizado de las clases de la imagen reducirá el aspecto de píxeles aislados dentro de otras clases, este comportamiento es típico de los clasificadores orientados al análisis píxel a píxel. El resultado es que los píxeles aislados pasarán a formar parte de clases más grandes y la percepción de la imagen será la habitual sobre la cubierta terrestre. El filtrado *Majority* examina las clases de los píxeles vecinos [87]. Por lo tanto la información de los píxeles vecinos es explotada con este tipo de filtros, esta información se añade a la hora del postprocesado para generar nuestras nuevas clases.

3.3. Evaluación de calidad de la clasificación

Desde el punto visual tenemos el resultado de nuestra clasificación, una imagen categorizada en una serie de clases. Lo que vamos a realizar ahora es cuantificar su precisión.

Toda clasificación conlleva un cierto margen de error, en función de la claridad de los datos o de la rigurosidad del método empleado. Por ello, resulta obligado aplicar algún tipo de procedimiento de verificación, que permita cuantificar ese error y, en función de él, valorar la calidad final del trabajo y su aplicabilidad operativa.

Para realizar la estimación de la exactitud alcanzada por nuestra imagen resultante se puede hacer por diversos criterios, nos centraremos en dos:

- i. Comparando nuestro inventario de clasificación con el obtenido por otras fuentes convencionales: el método más sencillo para estimar la precisión conseguida con nuestro mapa, se basa en calcular las diferencias entre el inventario ofrecido por la clasificación y por otras fuentes que se consideren fiables (estadísticas oficiales, cartografía de detalle, etc).
- ii. Seleccionando unas áreas de verificación para las que se cuenta con la cubierta realmente presente en el suelo: consiste en seleccionar, con

posterioridad a la clasificación, una serie de áreas de test diferentes a las de entrenamiento, que permitan obtener una medida independiente del acuerdo entre mapa y verdad-terreno. Este procedimiento pasa por realizar un muestreo del área de estudio, a fin de obtener las medidas de campo necesarias para verificar los resultados de la clasificación.

En nuestro Trabajo Fin de Máster nos hemos centrado en el segundo criterio, realizando un nuevo conjunto de regiones de interés para verificar los resultados (ROIs de evaluación). Para realizar cada región tendremos que hacer un muestreo de nuestra imagen, dentro de cada una de las clases. Existen varias técnicas de muestreo para realizar nuestra región de test:

- i. Aleatorio simple. Los elementos a verificar se eligen de tal forma que todos cuenten con la misma probabilidad de ser seleccionados, y que la elección de uno no influya en la del siguiente.
- ii. Aleatorio estratificado. La muestra se realiza dividiendo la población en regiones o estratos, de acuerdo a una variable auxiliar.
- iii. Sistemático. La muestra se distribuye a intervalos regulares, a partir de un punto de origen señalado aleatoriamente.
- iv. Sistemático no alineado. Modifica el esquema anterior, gracias a variar aleatoriamente -en cada línea y columna del muestreo- una coordenada, manteniendo fija la otra.
- v. Por conglomerados. Se selecciona como unidad de muestra un grupo de observaciones, denominado conglomerado (cluster), en lugar de individuos aislados. En otras palabras, en cada punto a verificar se toman varias muestras, en puntos vecinos al elegido aleatoriamente, y de acuerdo a un esquema.

En la Figura 3.9 podemos apreciar cada una de las técnicas de muestreo explicadas:

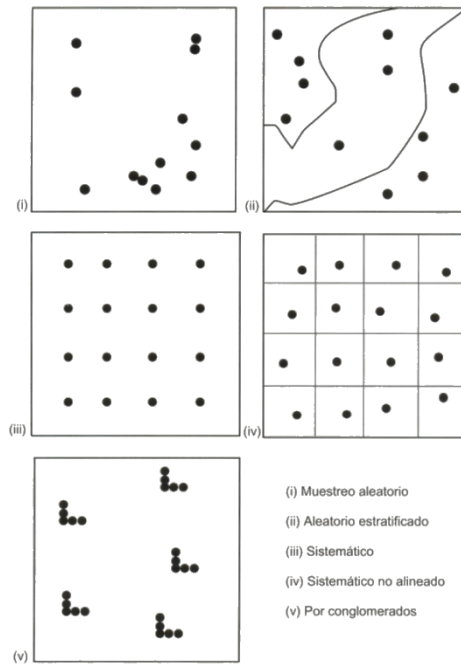


Figura 3.9. Técnicas de muestreo [57].

Ya elegida la región con la que vamos a verificar nuestros resultados pasaremos a hablar de la forma en que se nos presentarán nuestros resultados.

Consecuencia de la fase de muestreo será un listado de puntos de test, para los que poseemos tanto su cobertura real como la deducida por la clasificación. Con estos datos puede formarse una matriz, denominada de confusión puesto que recoge los conflictos que se presentan entre categorías. En nuestro caso, se generó una matriz para cada uno de los métodos de clasificación abordados en este proyecto. En la matriz de confusión las filas se ocupan por las clases deducidas de la clasificación, y las columnas por las clases de referencia. Lógicamente ambas tendrán el mismo número y significado; se trata, en resumen, de una matriz cuadrada: $n \times n$, donde n indica el número de categorías.

3.4. Técnica *Linear Spectral Unmixing* (LSU)

La técnica *Linear Spectral Unmixing* se emplea para determinar la abundancia relativa de los materiales que se representan en imágenes multiespectrales o hiperespectral basándose en las características espectrales de los materiales. La reflectancia en cada píxel de la imagen es una combinación lineal de la reflectancia de cada material (o *endmember*) presente dentro del píxel [88].

Por ejemplo, si el 25% de un píxel contiene material A, 25% del píxel contiene

material B, y 50% del píxel contiene material C, el espectro para ese píxel es un promedio ponderado de 0,25 veces el espectro de material A, más 0,25 veces el espectro de material B, más 0,5 veces el espectro de material de C. Por lo tanto, dado el espectro resultante (los datos de entrada) y los espectros *endmember*, el método *Linear Spectral Unmixing* resuelve los valores de abundancia de cada *endmember* para cada píxel.

En la Figura 3.10 podemos observar una imagen multispectral de 12 bandas de la cual se han seleccionado tres píxeles. El primero, sería un píxel mezcla entre suelo y rocas. El segundo, un píxel puro de agua y en tercer lugar un píxel mezcla de vegetación y suelo. A la derecha podemos apreciar las firmas espectrales de cada uno de los píxeles extraídos. Gracias a la técnica *Linear Spectral Unmixing*, seríamos capaces de obtener un mapa de abundancias en el que cada píxel estaría formado por la parte proporcional de cada clase hallada en ese píxel. Gracias a ese mapa de abundancias, estaremos en disposición de generar la cartografía de interés para este Trabajo Fin de Máster.

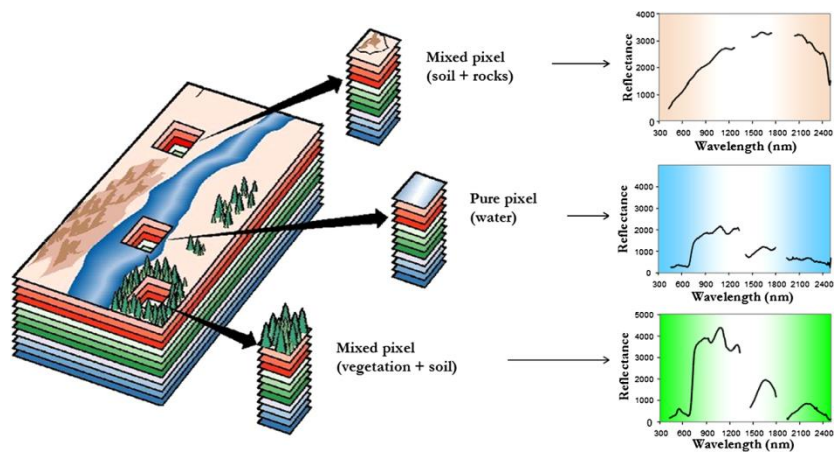


Figura 3.10. Imagen multispectral con píxeles mezclados y sus firmas espectrales [89].

El número de *endmembers* debe ser menor que el número de bandas espectrales, y todos los *endmembers* en la imagen deben ser utilizados. Los resultados de esta técnica son altamente dependientes de los *endmembers* de entrada, es decir, el cambio de *endmembers* cambiará directamente los resultados.

3.5. Resumen

En este capítulo se han explicado diversas técnicas de procesamiento de imágenes de satélite, tomando especial interés la clasificación supervisada.

Se ha expuesto, en primer lugar, la metodología a emplear para aplicar la

clasificación supervisada a las imágenes satelitales. Dicha metodología cuenta con tres fases principales: entrenamiento, asignación y post-procesado. La fase de entrenamiento es de vital importancia ya que deben quedar bien definidas las clases de interés del trabajo a realizar. Para ello, se culmina dicha fase con un estudio de separabilidad entre clases. Llegados a este punto pasaríamos a la fase de asignación, fase en la cual se asignaría cada píxel de la imagen a una de las clases definidas en la etapa anterior. Para realizar dicha asignación, se explican dos clasificadores supervisados: *Maximum Likelihood* y *Support Vector Machine*. Posteriormente llegaríamos a la fase de post-procesado. Dicha fase se encarga de suavizar los resultados obtenidos tras la fase de asignación con un filtrado *Majority*. Como etapa final de esta metodología, se explican los conceptos necesarios para realizar una evaluación de calidad de los resultados obtenidos.

Finalmente se explica la técnica de procesado de imágenes *Linear Spectral Unmixing*, que será empleada brevemente en este Trabajo Fin de Máster.

Capítulo 4: Generación de cartografía de vegetación y sustrato en Maspalomas

4.1. Datos y clases de interés

En este apartado detallaremos las características y especificaciones de la imagen a emplear en el Trabajo Fin de Máster que venimos desarrollando. Posteriormente, desglosaremos las clases que pretendemos discriminar en la Reserva Natural de las Dunas de Maspalomas. Además, se analizarán las medidas de campo y los instrumentos empleados para su obtención. Por último, expondremos las cartografías existentes de la zona de Maspalomas (Gran Canaria) y determinaremos si, para la evaluación de nuestros resultados, será de utilidad el empleo de las mismas.

4.1.1. Imagen WorldView 2

En este Trabajo Fin de Máster, vamos a emplear una imagen multiespectral que ha sido capturada por el satélite WorldView 2. Dicho satélite nos proporciona la imagen principal con 8 bandas multiespectrales y una imagen pancromática de una única banda pero con mejor resolución espacial que la imagen multiespectral. Gracias al proyecto ARTeMISat (CGL2013-46674-R), financiado por el Ministerio de Economía y Competitividad de España, estamos en disposición de emplear en este trabajo la imagen de la Figura 4.1 que se corresponde con una imagen multiespectral del Sur de Gran Canaria tomada el 4 de Junio de 2015. El área de interés para este trabajo será el correspondiente a la Reserva Natural de las Dunas de Maspalomas cuya extensión es apreciable en la Figura 4.2.

Debemos destacar que las Figuras 4.1 y 4.2 han sido pre-procesadas por lo que son el resultado de aplicar las adecuadas correcciones radiométricas y atmosféricas sobre la imagen satelital. Dichas correcciones ya han sido explicadas en capítulos anteriores por lo que no entraremos en detalle.

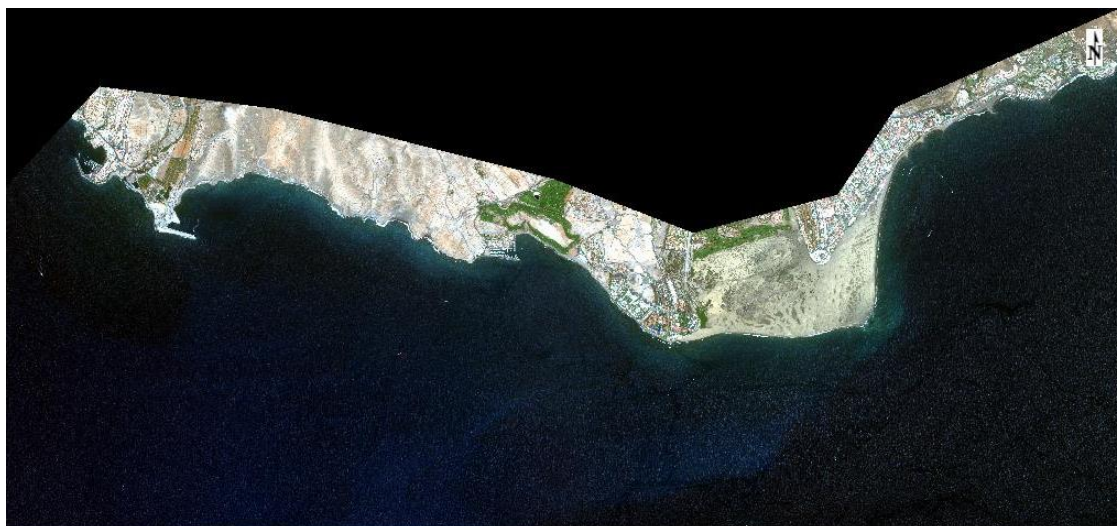


Figura 4.1. Imagen multiespectral del Sur de Gran Canaria tomada en Junio de 2015 [90].



Figura 4.2. Zoom de la imagen multiespectral en la zona de la Reserva Natural de las Dunas de Maspalomas.

4.1.2. Clases de interés

Una vez conocida la imagen y la zona sobre la que vamos a trabajar, debemos estudiar las clases de especial interés que se pretenden diferenciar en este Trabajo Fin de Máster. Dichas clases pueden ser separadas en dos grandes grupos: vegetación y sustrato.

En el grupo de vegetación, diferenciaremos las clases:

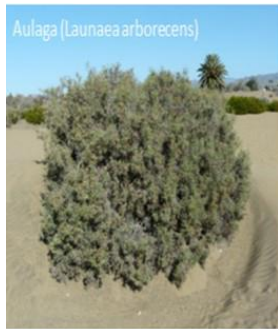
- **“*Tamarix canariensis*”**: más conocido como tarajal, taraje o taray de Canarias, es un pequeño árbol de la familia de las tamaricáceas.



El color empleado para distinguir esta especie en las cartografías ha sido, en su composición RGB, (0 128 0).



- **“*Launaea arborescens*”**: conocida como aulaga, cardaviejo o escarramoña entre otros, es una especie de fanerógama de la familia Asteraceae propia de las zonas costeras del sudeste de España, Islas Canarias y noroeste de África llegando por el sur hasta Mauritania.



El color empleado para las cartografías que vamos a generar es (196 196 0).



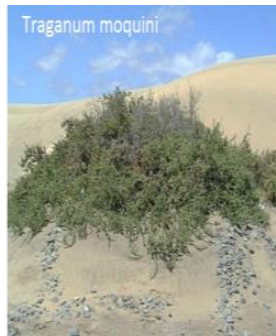
- **“*Juncus acutus*”**: El junco espinoso es una especie botánica de fanerógama de la familia de las juncáceas.



El color empleado para distinguir esta especie en los mapas generados será (0 220 200).



- **“*Traganum moquinii*”**: es una planta arbustiva nativa en las islas, que crece en zonas arenosas costeras. Esta especie se incluye en el Catálogo de



Especies Amenazadas de Canarias.

El color empleado para el “*Traganum moquinii*” será (110 190 160).



- **“*Suaeda mollis*”**: es una especie nativa en las islas. Se trata de un arbusto de 40-60 cm, con hojas subcilíndricas o globosas, de hasta 5 mm, pecioladas



y de color glauco a rojizo. Se conoce como "matomoro brusquillo".

El color que emplearemos para la “*Suaeda mollis*” es (128 128 192).



- **“*Tetraena fontanesii*”**: conocida como “*Zygophyllum fontanesii*”, también llamada Uva de mar, Uvilla o Salado baboso es una especie de arbusto









suculento autóctono de ciertos archipiélagos de la Macaronesia y del Noroeste de África. Está presente en todas las islas Canarias excepto La Palma.

El color empleado es el (180 200 250).



Por otro lado, en el grupo de sustrato distinguiremos:

- **Arena seca**: es la mayor parte de la arena que forma las Dunas. El color empleado para generar las cartografías será (240 240 140). 
- **Arena húmeda**: zonas sombrías y más frescas de la arena de las Dunas. Distinguiremos la arena húmeda con el color (205 205 177). 
- **Costra salina**: formada como restos que deja el mar tras bañar la arena. Emplearemos el blanco para diferenciarla (255 255 255). 

- **Charca:** abarcará exclusivamente a la Charca de Maspalomas. El color para distinguirla será (0 100 200). 
- **Aluviales:** son suelos de origen fluvial, poco evolucionados aunque profundos. Se forman de materiales transportados por corrientes de agua. Utilizaremos el color (196 127 0). 
- **Callaos:** piedras pequeñas y redondeadas por la erosión del tiempo. El color que representará a este sustrato en los mapas será (145 145 145). 

A modo de resumen, hemos añadido en la Tabla 4.1 la leyenda de colores que vamos a emplear en las cartografías generadas para este Trabajo Fin de Máster.








Vegetación	Gama de colores empleada	Sustrato	Gama de colores empleada
<i>“Tamarix canariensis”</i>		Arena seca	
<i>“Launaea arborescens”</i>		Arena húmeda	
<i>“Juncus acutus”</i>		Costra salina	
<i>“Traganum moquinii”</i>		Charca	
<i>“Suaeda mollis”</i>		Aluviales	
<i>“Tetraena fontanesii”</i>		Callaos	

Tabla 4.1. Leyenda de colores para las clases a distinguir.

4.1.3. Medidas de campo

Dentro del proyecto ARTeMISat se ha llevado a cabo una campaña para la toma de medidas de campo de las distintas especies y sustratos que conviven en las Dunas de Maspalomas. Gracias a dicha campaña y a la colaboración de expertos en la materia, cada una de las clases que se pretenden distinguir en este Trabajo Fin de Máster ha quedado identificada y localizada. En la Figura 4.3 podemos apreciar varios momentos de la toma de medidas, especialmente para las especies *“Tamarix canariensis”*, *“Juncus acutus”* y *“Suaeda mollis”* en las Figuras 4.3 (a), (b) y (c) respectivamente y la costra salina en la Figura 4.3 (d).

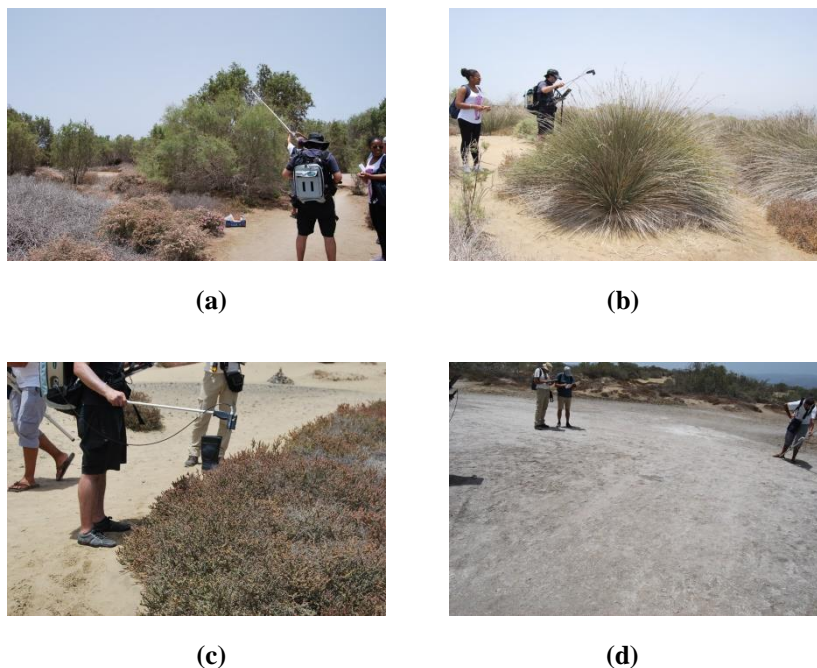


Figura 4.3. Toma de medidas de la campaña de campo sobre: (a) “*Tamarix canariensis*”, (b) “*Juncus acutus*”, (c) “*Suaeda mollis*” y (d) Costra salina.

La toma de medidas se ha llevado a cabo con un radiómetro, concretamente el de la Figura 4.4, siendo éste capaz de detectar y medir la intensidad de energía reflejada de cada especie o sustrato.



Figura 4.4. Radiómetro empleado en la campaña de campo.

La Figura 4.5 se corresponde con un ejemplo de medida empleando el radiómetro, concretamente con el momento de la captura de la respuesta espectral de la Charca de Maspalomas. Adicionalmente y a modo de ejemplo, hemos añadido en la Figura 4.6 la hoja de características y resultados tras estudiar la especie “*Juncus acutus*”. De esa forma, cada una de las especies del estudio queda identificada y diferenciada por su firma espectral media \pm la desviación estándar.



Figura 4.5. Empleo del radiómetro para medir la respuesta espectral de la Charca de Maspalomas.

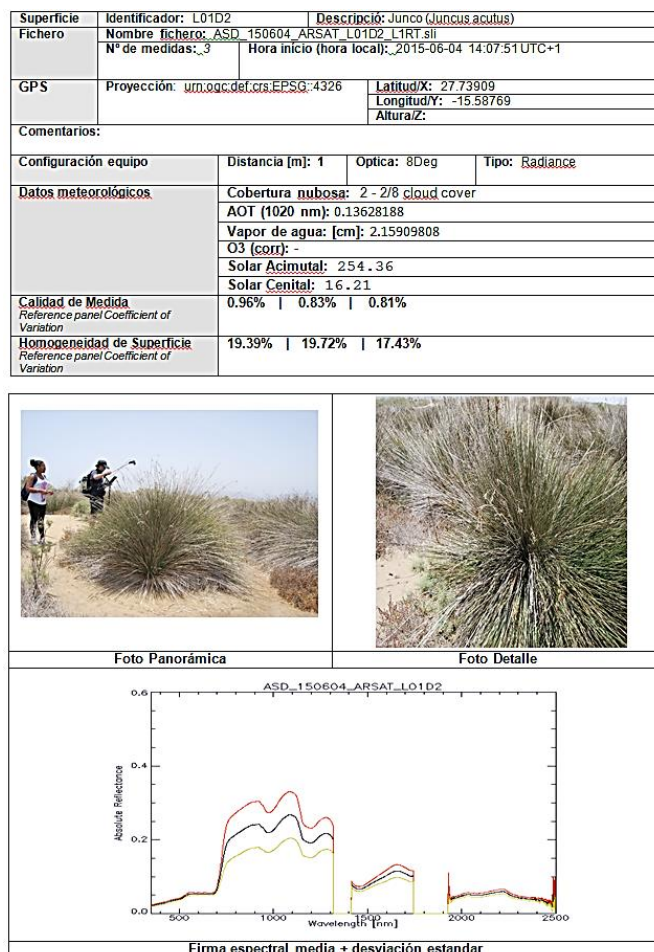


Figura 4.6. Hoja de características y resultados de la especie “*Juncus acutus*”.

Dada la planificación temporal estipulada para este Trabajo Fin de Máster, aprovecharemos de esta campaña de campo los mapas de localización de la Figura 4.7, quedando reservado para líneas futuras de trabajo el resto de los datos adquiridos. Gracias a estos mapas, podremos seleccionar las distintas especies vegetales y los tipos de sustrato para generar posteriormente las cartografías de interés.

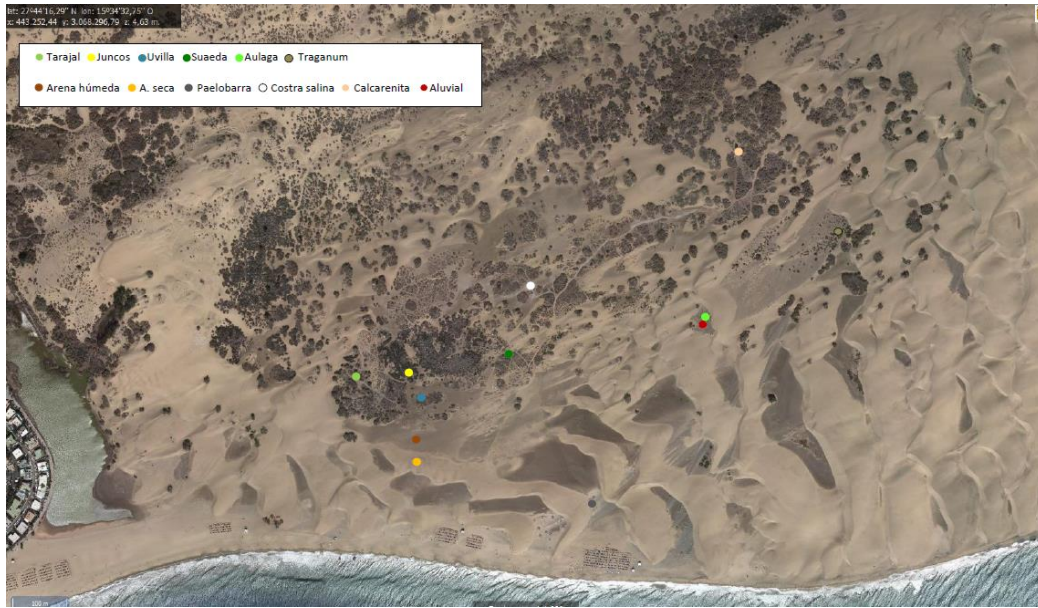


Figura 4.7. Mapa de localización de las especies vegetales y los distintos tipos de sustrato.

4.1.4. Cartografías de interés de Maspalomas

Motivados por el interés y la importancia del estudio y la conservación de la Reserva Natural de las Dunas de Maspalomas, debemos analizar la cartografía existente y aprovechar la información que nos puedan aportar para la evaluación de nuestros resultados.

Generar este tipo de cartografía es muy complejo y costoso y apenas se dispone de información. Además, los mapas existentes que identifican zonas de vegetación no son nada específicos. Si intentamos buscar información sobre cartografías de vegetación en la Reserva, nos encontraremos mapas como el de la Figura 4.8 del Gobierno de Canarias, actualizado por última vez con fecha 1999-2001. Además se aprecia que la cartografía obtenida del Gobierno de Canarias [91] ha sido un proceso de selección de áreas y no de especies de forma individualizada, sin importar que en esas zonas marcadas haya o no otro tipo de vegetaciones a parte de las que se mencionan. Si nos centramos en los tipos de sustrato, esta tarea resulta igualmente complicada puesto que no existen dichos mapas.

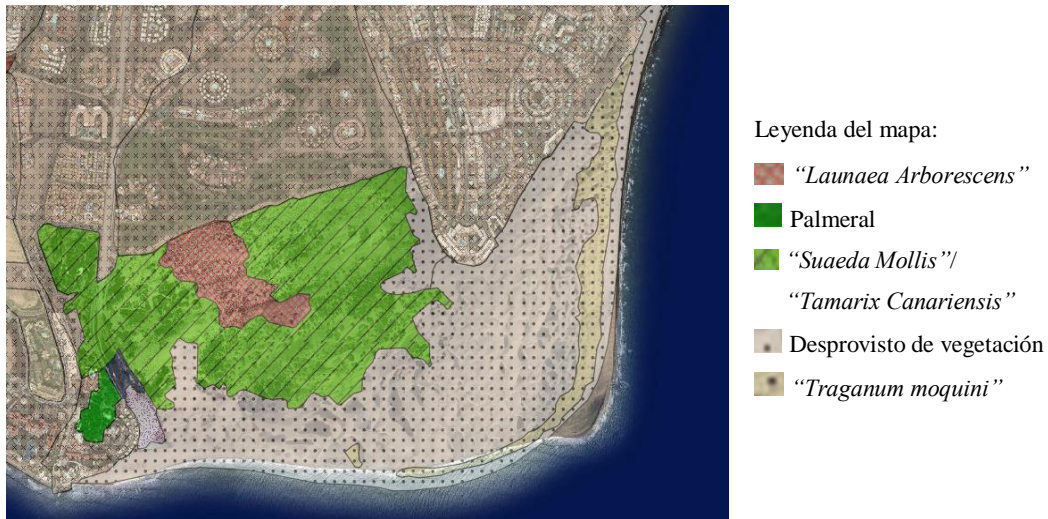


Figura 4.8. Mapa de vegetación facilitado por el Gobierno de Canarias [91].

En concordancia con el Sistema de Información sobre Ocupación del Suelo de España (SIOSE) [92], integrado dentro del Plan Nacional de Observación del Territorio (PNOT) cuyo objetivo es generar una base de datos de Ocupación del Suelo para toda España, deberían existir mapas de mayor información. Sin embargo, en la Figura 4.9 podemos observar el Mapa de Ocupación de Suelo, centrado en la Reserva, que nos ofrece el Gobierno de Canarias. Si solicitamos información en un punto de la Reserva Natural de las Dunas de Maspalomas, nos proporcionan los datos de la Figura 4.10. En ella podemos apreciar que la Reserva se incluye en el grupo de suelo desnudo y que las clases de ocupación son playas, dunas y arenales. Evidentemente en este mapa no se han tenido en cuenta, por un lado, las variadas especies vegetales que habitan en las Dunas y, por otro, los distintos tipos de sustrato que se entremezclan en las mismas.



Figura 4.9. Mapa de Ocupación del Suelo facilitado por el Gobierno de Canarias [91].

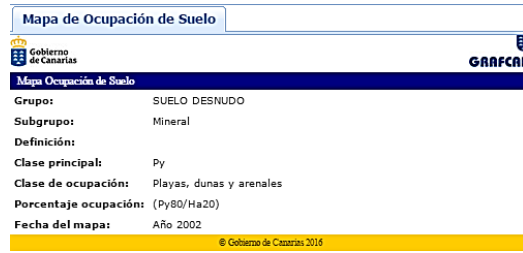


Figura 4.10. Información del Mapa de Ocupación de Suelo sobre la Reserva Natural de las Dunas de Maspalomas.

Las cartografías analizadas hasta este momento no nos permitirían evaluar los resultados de forma óptima por lo que nos apoyaremos en datos y mapas extraídos de la Tesis Doctoral de Don Antonio Ignacio Hernández [93]. En la Figura 4.11 se puede apreciar la cartografía que emplearemos como referencia para evaluar la vegetación de los mapas generados a lo largo de este Trabajo Fin de Máster.

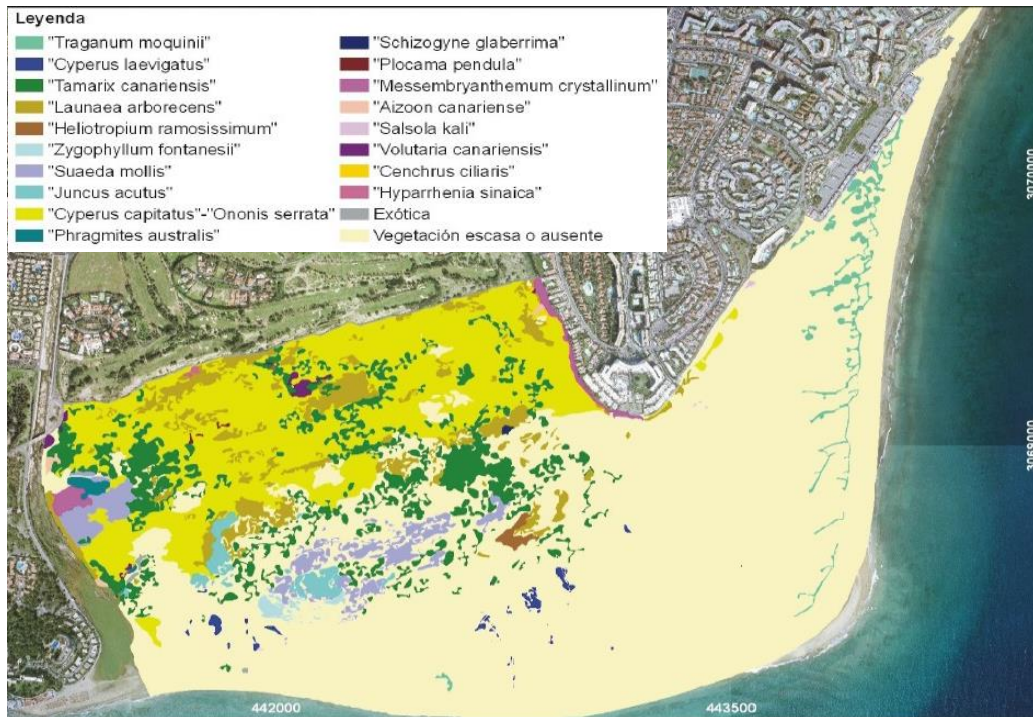


Figura 4.11. Cartografía de vegetación de la Reserva Natural de las Dunas de Maspalomas [93].

Adicionalmente, se hará uso de la cartografía a nivel de sustrato generada en la Tesis Doctoral de Don Luis F. Hernández Calvento [94]. Este mapa, observable en la Figura 4.12, facilitará la tarea de evaluación de sustratos en las cartografías generadas a lo largo de este trabajo. Concretamente, aprovecharemos la información de la localización de los depósitos aluviales y los cordones de cantos o callaos.

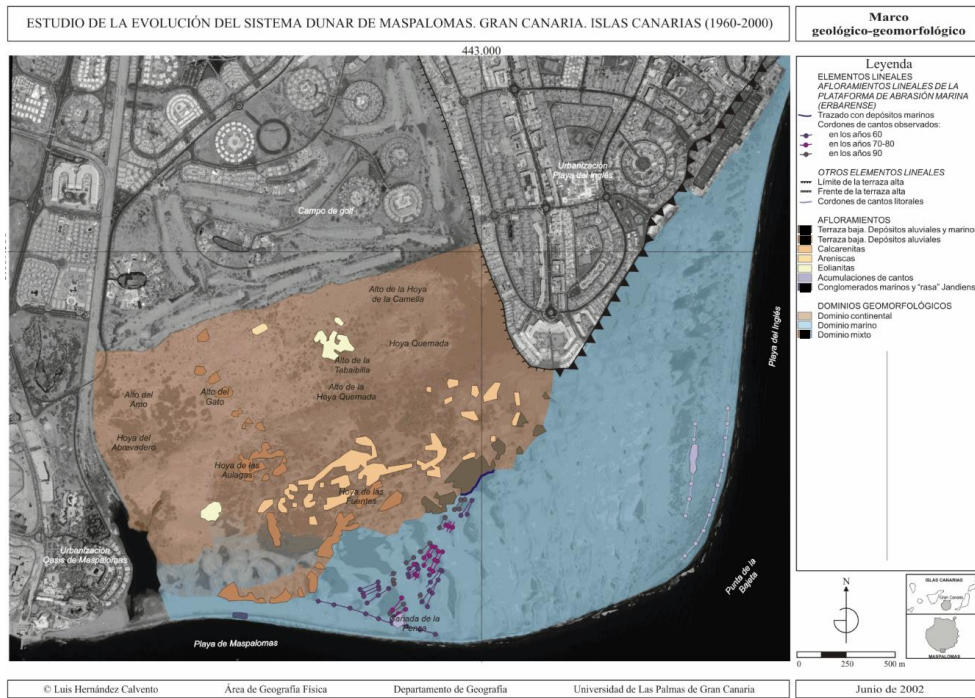


Figura 4.12. Cartografía de afloramientos y elementos lineales en la Reserva Natural de las Dunas de Maspalomas [94].

4.2. Metodología

En este apartado, definiremos paso a paso la metodología a seguir para desarrollar nuestro Trabajo Fin de Máster. Para ello, nos apoyaremos en la Figura 4.13.

Lo primero que debemos hacer es tener a nuestra disposición las imágenes multiespectral y pancromática de la zona de estudio de este trabajo. En segundo lugar, debemos aplicarle a dichas imágenes las correcciones radiométricas y atmosféricas explicadas en capítulos anteriores.

A partir de este momento, tendremos dos líneas de trabajo: por un lado la imagen multiespectral y por el otro la imagen fusionada, siendo ésta el resultado de aplicar la Técnica de Fusión Gram-Schmidt (*pansharpening*) a las imágenes multiespectral y pancromática.

Para poder avanzar, es necesario desarrollar la parte del esquemático remarcada en color rojo. Esta tarea es de vital importancia para augurar unos resultados óptimos tras finalizar la ejecución total del esquema. Lo primero que debemos hacer es definir las clases que queremos identificar y exponer la leyenda de colores que diferencie a cada una de ellas. Una vez definidas las clases, debemos generar dos regiones de interés

(ROIs): una de entrenamiento y otra de evaluación. Dichas ROIs no pueden tener píxeles en común puesto que ello supondría la obtención de unos resultados falseados. Generadas las ROIs, estudiaremos la separabilidad de las clases hasta que su valor sea el adecuado para enfrentarnos a la clasificación.

Una vez finalizado el proceso de creación y estudio de las ROIs, tanto la imagen multiespectral como la imagen fusionada deberán ser enmascaradas, reduciendo así el tamaño de la imagen y, en consecuencia, los tiempos de cómputo. Dicha máscara nos permitirá analizar de forma aislada la Reserva Natural de las Dunas de Maspalomas.

Llegados a este punto pasaremos a la fase de asignación, que se corresponde con la clasificación de las imágenes multiespectral y fusionada, ambas enmascaradas, empleando la ROI de entrenamiento. Adicionalmente aplicaremos técnicas de *unmixing*, utilizando para ello la imagen multiespectral y la ROI de entrenamiento.

De forma opcional, hemos añadido una etapa de post-clasificación. Dicha etapa nos permitirá suavizar los resultados de la clasificación empleando un tipo de filtrado denominado *Majority*.

Todos los resultados obtenidos, pasarán a la etapa de evaluación. Esta etapa nos proporcionará las matrices de confusión y los coeficientes Kappa a partir de la utilización de la ROI de evaluación y las cartografías resultantes de la etapa de clasificación con/sin *Majority*.

Evaluadas las cartografías obtenidas, pasaremos a la etapa final en la que se analizarán los resultados obtenidos y se hará una comparativa con el mapa de referencia [93]. De esa forma, podremos analizar las cartografías de nuestro estudio frente a la información real del terreno.

Para realizar todo el proceso explicado y expuesto en la Figura 4.13, separaremos en un apartado los resultados cartográficos empleando la imagen multiespectral y, en otro, los resultados tras utilizar la imagen fusionada. Dicha distinción nos permitirá analizar los resultados cómodamente.

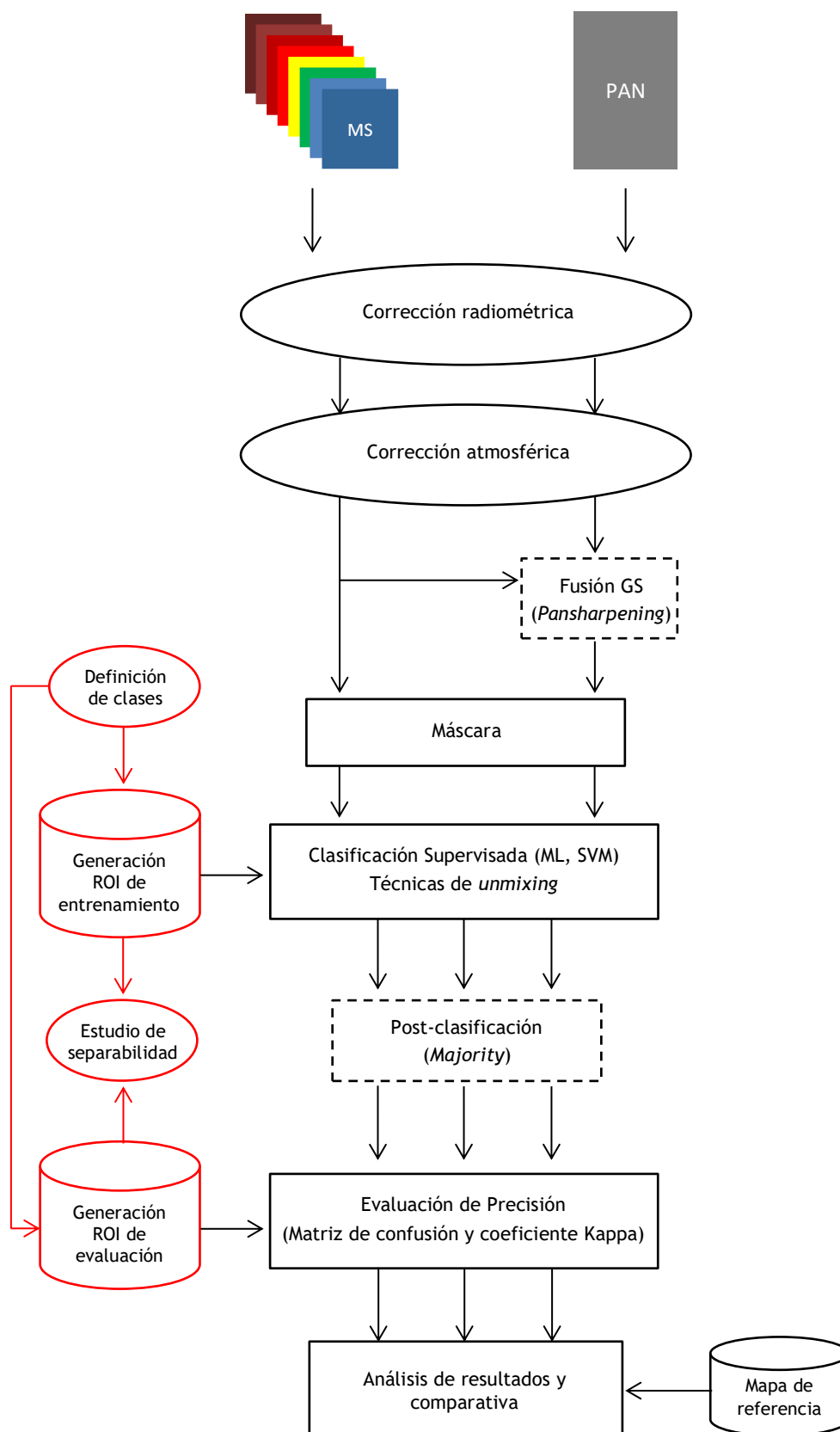


Figura 4.13. Metodología para la obtención de cartografías.

4.3. Software ENVI

La creciente dependencia de las imágenes geospaciales hace cada vez más importante la capacidad de obtener la información necesaria a partir de estos datos. Es esencial disponer de herramientas y aplicaciones que ayuden a extraer información de forma sencilla y precisa independientemente de que nuestro foco esté en la inteligencia, la investigación científica o la planificación y toma decisiones.

Lanzado originariamente en 1994, ENVI (*Environment for Visualizing Images*), es un moderno sistema de procesamiento de imágenes diseñado para proporcionar análisis de los datos obtenidos por teledetección desde aviones y satélites. Proporciona un entorno potente, novedoso y de fácil uso para presentar y analizar imágenes de cualquier tamaño y tipo de datos en un amplio rango de plataformas [95].

La última versión de ENVI hace más eficiente que nunca el proceso de trabajo analítico permitiendo tener información de una forma más rápida. Además, nos proporciona una interfaz de usuario más intuitiva que facilita las distintas tareas a desarrollar.

ENVI es conocido por su amplio conjunto de herramientas de análisis de imágenes y por sus flujos de trabajo automatizados, que permiten extraer información crítica de las imágenes.

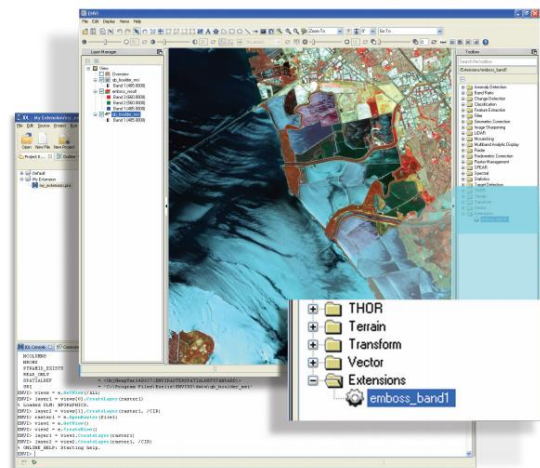


Figura 4.14. Interfaz de usuario de ENVI 5.0 [96].

Entre otras herramientas, ENVI nos ofrece un amplio rango de posibilidades de análisis espectral y análisis de datos con las que podemos realizar, entre otras, estadísticas geospaciales tales como autocorrelación y semi-varianza, extracción de características lineales, detección de cambios, medición de rasgos, modelado de características topográficas, detección e identificación de objetivos, clasificación de imágenes usando métodos supervisados y no supervisados, identificación de firmas espectrales usando librerías espectrales completas e identificar características de interés.

ENVI está completamente escrito en IDL (*Interactive Data Language*), Lenguaje de Datos Interactivo. IDL es un potente lenguaje de programación estructurado, basado en matrices, que proporciona un procesamiento de imágenes integrado, grandes

capacidades de visualización y herramientas GUI fáciles de usar. Se requiere IDL para ejecutar ENVI, y ENVI recibe gran parte de su flexibilidad de la potencia de IDL.

En definitiva, el software que emplearemos para realizar nuestro Trabajo Fin de Máster será el ENVI 5.0 debido a sus innumerables prestaciones.

4.4. Resultados para la imagen sin fusionar

Detallada anteriormente la metodología a seguir y asumiendo que nuestras imágenes multiespectral y pancromática ya han sido corregidas radiométrica y atmosféricamente, vamos a exponer independientemente los resultados de la imagen multiespectral y la fusionada. Comenzaremos con la imagen multiespectral sin fusionar.

4.4.1. Regiones de interés (ROIs) y cálculo de la separabilidad

Como ya hemos explicado, una región de interés se realiza definiendo, en nuestra imagen a clasificar, una serie de píxeles característicos de cada clase. Estos píxeles pueden estar definidos con puntos aislados o formando figuras geométricas.

En apartados anteriores de este Trabajo Fin de Máster hemos explicado detalladamente cómo se realizan estas regiones, por lo que, a modo de ejemplo observaremos únicamente la primera región definida. A continuación pasaremos directamente a mostrar los resultados de separabilidad obtenidos para las regiones de interés (ROIs) generadas.

4.4.1.1. ROI de entrenamiento para la imagen original

La primera ROI definida, denominada ROI de entrenamiento, se empleará para entrenar a nuestro clasificador. Dicha ROI está compuesta de píxeles puros y mezclados. Estos píxeles mezclados son píxeles correctos en los que se puede apreciar mayoritariamente la clase a la que pertenecen pero no llegan a ser píxeles puros.

Para este caso, el número de píxeles por clase varía entre 25 y 500 con la salvedad de las clases Arena Seca y Arena Húmeda que contienen entre 1000 y 3000 píxeles por ser clases más abundantes. En la Figura 4.15 (a), tenemos la imagen completa de estudio. En este caso y para este estudio en concreto, realizaremos un recorte y nos centraremos en estudiar la zona de la Reserva Natural de las Dunas de Maspalomas. Dicho recorte se puede observar en la Figura 4.15 (b).

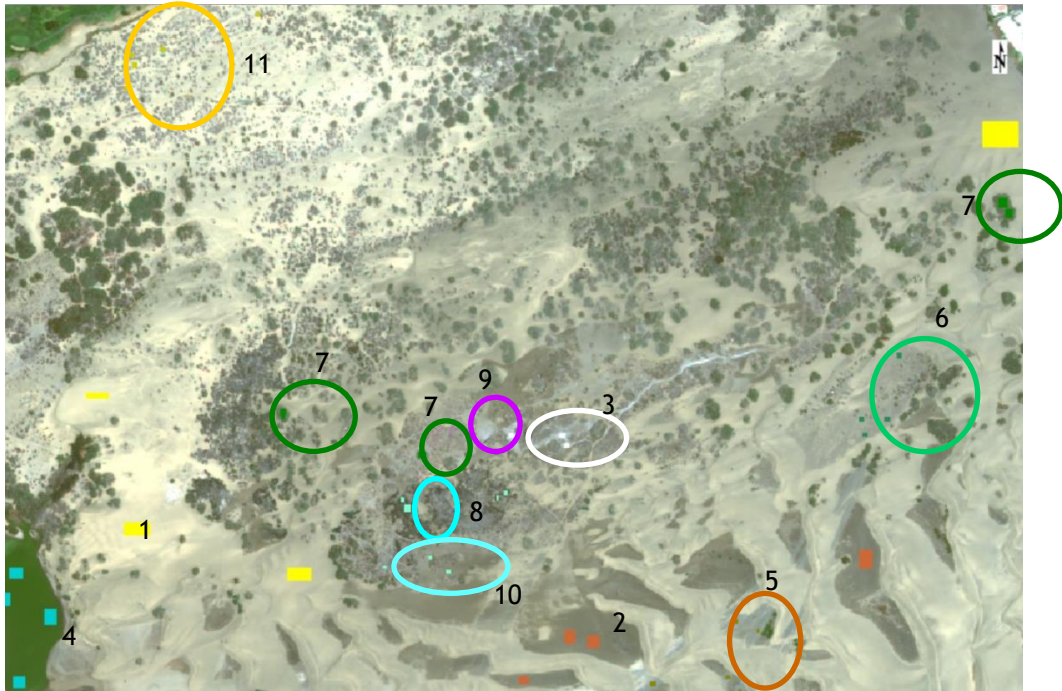
Si realizamos un zoom en la zona de más vegetación de la Reserva, podemos observar en la Figura 4.15 (c) las zonas que nos definen las distintas clases a discriminar. Para facilitar la lectura de la imagen, hemos añadido un número a cada clase. A nivel de sustrato, queremos diferenciar Arena seca (1), Arena húmeda (2), Costra Salina (3), Charca (4), Callaos (5) y Aluviales (6). En cuanto a la vegetación, diferenciaremos las clases “*Tamarix canariensis*” (7), “*Juncus acutus*” (8), “*Suaeda mollis*” (9), “*Tetraena fontanesii*” (10), “*Launaea arborescens*” (11) y finalmente “*Traganum moquinii*” (12). Esta última clase del grupo de vegetación aparece en la Figura 4.15 (b) en el zoom remarcado (naranja) al Este de la Reserva Natural de las Dunas de Maspalomas puesto que es la zona en la que mayoritariamente habita dicha especie.



(a)



(b)



(c)

Figura 4.15. ROI 1 de entrenamiento: (a) Imagen completa, (b) Zoom en la Reserva Natural de las Dunas de Maspalomas, y (c) Zonas de entrenamiento para la parte central de la Reserva.

De esta forma ya hemos definido mediante la selección de píxeles cada una de las clases a discriminar y, por tanto, tenemos generada nuestras regiones de interés de entrenamiento (ROI de entrenamiento). Posteriormente se creará la ROI de evaluación.

A continuación hemos obtenido, siguiendo el procedimiento explicado en capítulos anteriores, el fichero que nos facilita información relativa a la separabilidad entre clases. Dicha información ha sido dividida en dos partes. La primera parte, formada por los índices de Jeffries-Matusita (JM) y la divergencia transformada que podremos encontrar en el Anexo1. La segunda, se corresponde con los pares de separabilidad obtenidos entre clases. Para una lectura comprensiva de los resultados, hemos plasmado los pares de separabilidad en la siguiente tabla:

CLASES	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales	<i>Tamarix Canariensis</i>	<i>Tragacanthum Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborecens</i>
Arena Seca	1,7969	1,9787	2,0000	1,7670	1,6919	1,9994	2,0000	2,0000	1,9999	1,8751	1,9996
Arena Húmeda		1,9977	2,0000	1,8896	1,9180	1,9989	2,0000	2,0000	2,0000	1,7312	2,0000
Costra Salina			2,0000	1,9974	1,9888	1,9999	2,0000	2,0000	2,0000	1,9950	2,0000
Charca				2,0000	2,0000	2,0000	2,0000	2,0000	2,0000	2,0000	2,0000
Callaos					1,9350	1,9997	2,0000	2,0000	2,0000	1,9275	2,0000
Aluviales						1,9999	2,0000	2,0000	2,0000	1,9513	2,0000
<i>Tamarix Canariensis</i>							1,7897	1,9658	1,9115	1,8939	1,9898
<i>Tragacanthum Moquinii</i>								1,9992	1,9162	1,8184	1,9259
<i>Juncus Acutus</i>									1,9983	1,9872	1,9999
<i>Suaeda Mollis</i>										1,3554	1,9860
<i>Tetraena Fontanesii</i>											1,9749

Tabla 4.2. Pares de separabilidad para la ROI de entrenamiento (distancia JM).

Basándonos en los resultados obtenidos encontramos que el peor par de separabilidad, resultado remarcado en rojo, se da entre las clases “*Suaeda Mollis*” y “*Tetraena Fontanesii*”. Por otro lado, observamos en verde los mejores pares de separabilidad alcanzando el valor óptimo. Analizando la Tabla 4.2, apreciamos que la clase Charca es totalmente separable del resto de las clases en estudio. Adicionalmente, apreciamos que las clases pertenecientes a sustrato se distinguirán con facilidad de las clases pertenecientes a vegetación. Como consecuencia de estos resultados podemos afirmar que habrá mayor probabilidad de error entre las clases “*Suaeda Mollis*” y “*Tetraena Fontanesii*”, pero realmente el resultado del resto de clases es bastante alto por lo que a priori la clasificación será viable para categorizar los tipos de vegetación y sustratos seleccionados.

Para poder apreciar el resultado de los pares de separabilidad, hemos reflejado los valores en la Figura 4.16. En ella podemos apreciar, por un lado, el mínimo par de separabilidad obtenido y, por otro, que el resto de pares de separabilidad superan el valor mínimo recomendado (1.7) de cara a obtener unos buenos resultados.

Para leer de forma sencilla la Figura 4.16 hemos añadido en la Tabla 4.3 una leyenda con los acrónimos empleados.

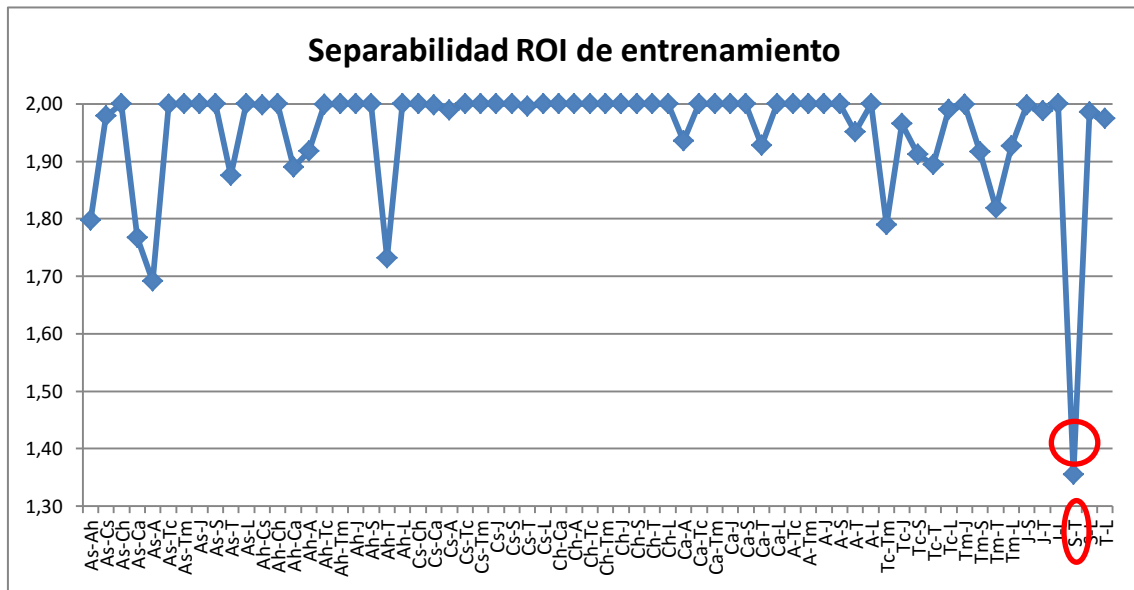


Figura 4.16. Representación de los pares de separabilidad para la ROI de entrenamiento (distancia JM).

Sustrato	Acrónimos	Vegetación	Acrónimos
Arena Seca	As	" <i>Tamarix canariensis</i> "	Tc
Arena Húmeda	Ah	" <i>Tragnum moquinii</i> "	Tm
Costra Salina	Cs	" <i>Juncus acutus</i> "	J
Charca	Ch	" <i>Suaeda mollis</i> "	S
Callaos	Ca	" <i>Tetraena fontanesii</i> "	T
Aluviales	A	" <i>Launaea arborescens</i> "	L

Tabla 4.3. Leyenda para la Figura 4.16.

Una vez explicada y desglosada la ROI de entrenamiento, aplicaremos el siguiente procedimiento a la ROI de evaluación: en primer lugar debemos detallar su composición a nivel de píxel. Además adjuntaremos en el Anexo 1 los resultados relativos a los índices de Jeffries-Matusita y la divergencia transformada. Por último se incluirá una tabla con los pares de separabilidad obtenidos y su respectiva explicación.

4.4.1.2. ROI de evaluación para la imagen original

Llegados a este punto, vamos a definir la región de interés de evaluación con la que mediremos la fiabilidad de los clasificadores tras haberlos entrenado con la región de interés de entrenamiento. Aplicando esta nueva ROI a las clasificaciones que realizaremos con la ROI de entrenamiento y obteniendo las matrices de confusión, podremos saber la exactitud de la clasificación. Por ese motivo, la ROI de evaluación no debe tener píxeles en común con la ROI de entrenamiento puesto que en ese caso estaríamos falseando los resultados.

A continuación definiremos una ROI de evaluación, que estará compuesta de píxeles puros y píxeles mezclados.

Hemos definido a cada una de las clases incluyendo un número de píxeles entre 25 y 500 con la salvedad de las clases Arena Seca y Arena Húmeda, que se definen en un rango de 1000 a 3000 píxeles. Una vez definidas las clases hemos calculado sus parámetros de separabilidad. Al igual que para la ROI de entrenamiento, los índices de Jeffries-Matusita y de la divergencia transformada para esta ROI de evaluación pueden ser consultados en el Anexo 1. En cuanto a los pares de separabilidad, se presentan en la Tabla 4.4.

En primer lugar, podemos apreciar que se da el peor par de separabilidad entre las clases “*Suaeda Mollis*” y “*Tetraena Fontanesii*” con un valor de 1.6905. Este par de clases coincide con el peor par obtenido en la ROI de entrenamiento. La diferencia recae en que la distancia JM es algo superior.

Adicionalmente, hemos generado la Figura 4.17 con cada uno de los pares de separabilidad obtenidos en la ROI de evaluación para facilitar la apreciación descrita anteriormente. Si nos fijamos en los valores de separabilidad de la Figura 4.17, vemos como todos superan ese límite de 1.7 para esperar un resultado bueno. Sin embargo, hemos remarcado en rojo el único par que queda fuera de ese umbral pero por una diferencia mínima, por tanto, auguramos unos resultados fiables a la hora de evaluar las clasificaciones realizadas previamente.

CLASES	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales	<i>Tamarix Canariensis</i>	<i>Tragnum Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborecens</i>
Arena Seca	1.8548	1.9557	2.0000	1.7445	1.7656	1.9999	1.9999	2.0000	1.9752	1.9961	1.9999
Arena Húmeda		1.9950	2.0000	1.8417	1.8734	1.9999	1.9999	2.0000	1.9923	1.9762	1.9999
Costra Salina			2.0000	1.9962	1.9789	1.9999	1.9999	2.0000	1.9951	1.9990	1.9999
Charca				2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000
Callaos					1.9327	2.0000	1.9999	2.0000	1.9980	1.9834	2.0000
Aluviales						2.0000	1.9997	2.0000	1.9751	1.9745	1.9999
<i>Tamarix Canariensis</i>							1.8758	1.9750	1.9989	1.9885	1.9983
<i>Tragnum Moquinii</i>								1.9999	1.8622	1.7850	1.8891
<i>Juncus Acutus</i>									1.9999	1.9994	1.9994
<i>Suaeda Mollis</i>										1.6905	1.9862
<i>Tetraena Fontanesii</i>											1.9707

Tabla 4.4. Pares de separabilidad para la ROI de evaluación (distancia JM).

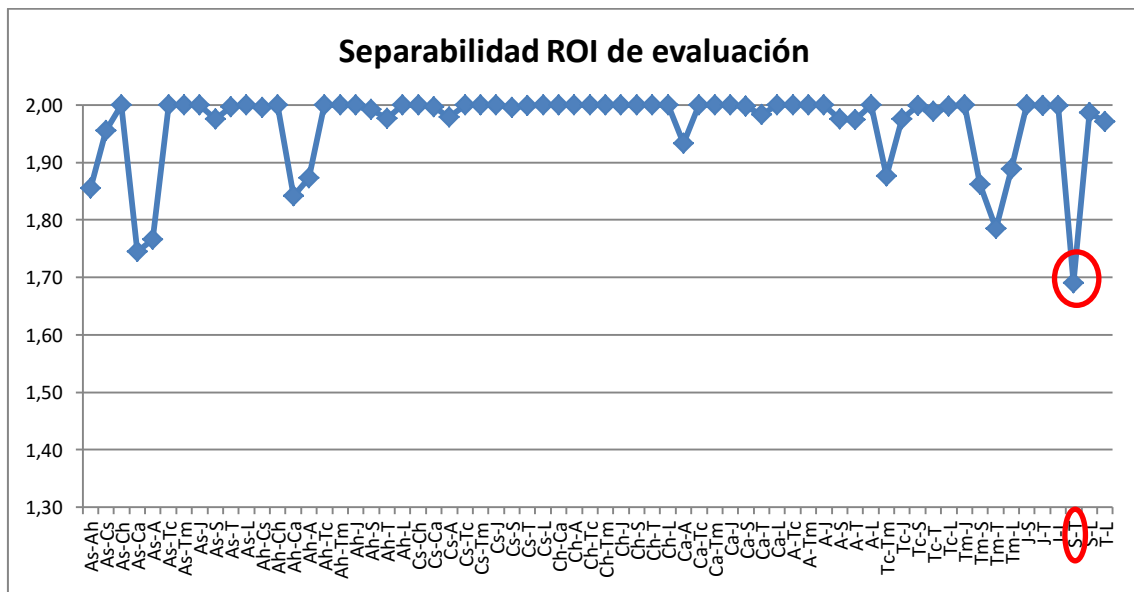


Figura 4.17. Representación de los pares de separabilidad para la ROI de evaluación (distancia JM).

Teniendo en cuenta los resultados obtenidos, hemos representado en la Figura 4.18 los pares de separabilidad, tanto para la ROI de entrenamiento como para la ROI de evaluación. De esta forma, podemos apreciar que ambas han quedado bien definidas con

La metodología empleada para construir la máscara es la que se describe a continuación. En primer lugar, debemos generar una ROI que delimite el área de interés. En segundo lugar, construimos la máscara con el comando *build mask* y ya tendremos la máscara preparada para aplicarla sobre nuestra imagen original.

4.4.3 Clasificación supervisada y matrices de confusión.

Llegados a este punto, ya disponemos de las regiones de interés y de la imagen enmascarada por lo que podemos empezar a clasificar La Reserva Natural de las Dunas de Maspalomas en la imagen de estudio.

Tal y como hemos comentado en capítulos anteriores, emplearemos 2 clasificadores supervisados: *Maximum Likelihood* y *Support Vector Machine*. En este primer estudio, los parámetros modificables de los clasificadores tendrán valores habituales por defecto y no serán ajustados. A continuación, realizaremos una comparativa entre los resultados obtenidos mediante la clasificación empleando la ROI de entrenamiento y cada uno de los clasificadores disponibles. Expondremos los resultados paso a paso de forma comparativa.

Adicionalmente, se ha generado un código con ayuda del software Matlab (versión 2015). Este código nos ha permitido generar los mapas finales de los resultados de la clasificación con los colores adecuados para cada clase a distinguir. Este trabajo ha sido necesario puesto que el software ENVI aplica ciertas modificaciones sobre los colores especificados y ha sido imposible de forma manual obtener los colores deseados. Mediante ese código, se recorre la imagen resultado que nos da ENVI tras la clasificación, se hace una transformación al cubo RGB de la imagen para poder trabajar en plano, se recorre píxel a píxel buscando los colores que ENVI les ha asignado y finalmente se les asigna el color deseado. El código completo se ha añadido en el Anexo 3.

Vamos a pasar a analizar los resultados obtenidos tras emplear la ROI de entrenamiento. En la Figura 4.20 podemos apreciar a simple vista, que hay ciertas diferencias en ambas clasificaciones realizadas si nos centramos en la parte central y superior de la región a estudio. Para poder extraer conclusiones de manera real y con datos estadísticos, vamos a evaluar nuestros resultados. Para ello aplicaremos la ROI de evaluación y obtendremos las matrices de confusión, el *overall accuracy* y el coeficiente Kappa para cada clasificación. Hay que indicar que la evaluación cualitativa

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

es importante y que la cuantitativa depende en cierta medida de la ROI de evaluación generada.

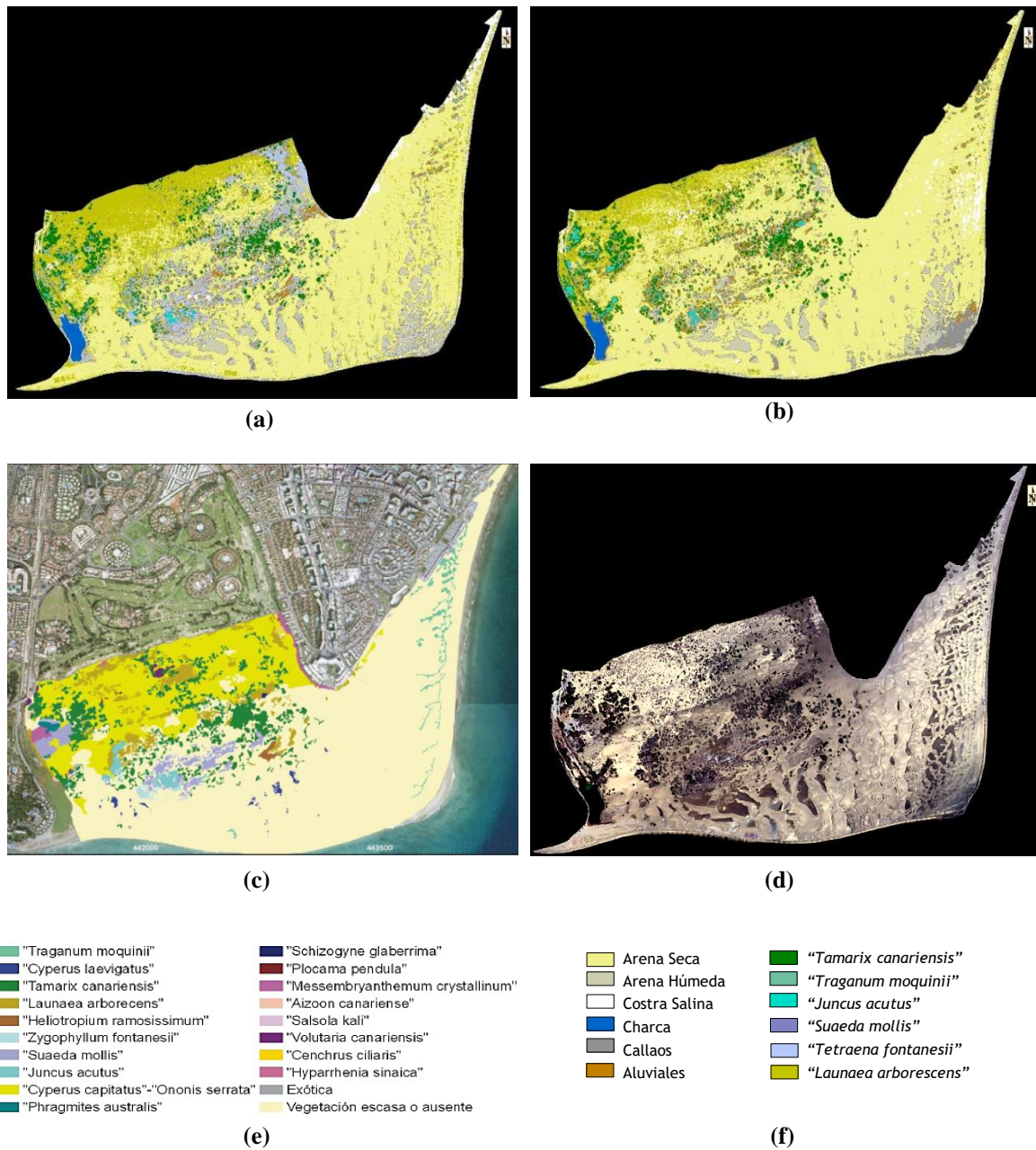


Figura 4.20. Clasificaciones con ROI de entrenamiento (a) *Maximum Likelihood*, (b) *Support Vector Machine*, (c) Mapa de referencia, (d) Imagen Original, (e) Leyenda del mapa de referencia, (f) Leyenda de los mapas de clasificación.

- **Evaluación con la ROI de evaluación sobre las clasificaciones realizadas**

En primer lugar vamos a analizar los resultados que se obtienen tras evaluar una imagen clasificada. En concreto, los resultados expuestos a continuación se corresponden con los obtenidos tras evaluar las clasificaciones *Maximum Likelihood* y

Support Vector Machine. En la Tabla 4.5 podemos observar el *overall accuracy* (precisión general) y el coeficiente Kappa que, como ya explicamos en capítulos anteriores, es el encargado de delimitar el grado de ajuste debido sólo a la exactitud de la clasificación prescindiendo del causado por factores aleatorios. En este caso hemos marcado en color verde el mejor resultado, que se corresponde al del clasificador *Maximum Likelihood*. Debemos recordar que en estos primeros resultados, los parámetros de los clasificadores han sido tomados por defecto.

Clasificador	<i>Overall Accuracy</i>	Coficiente Kappa (k)
<i>Maximum Likelihood</i>	91,5366%	0,8885
<i>Support Vector Machine</i>	87,5103%	0,8301

Tabla 4.5. *Overall Accuracy* y coeficiente Kappa. Clasificadores con parámetros por defecto.

Además del coeficiente Kappa y del *overall accuracy*, tras la evaluación se obtiene una matriz de confusión que nos permite evaluar los conflictos que se producen entre las clases. En la matriz de confusión las filas se corresponden con las clases deducidas de la clasificación y las columnas se corresponden con las clases de referencia. La diagonal representa los puntos en los que se ha producido concordancia entre las clases reales y las deducidas. Esta matriz es muy útil para medir la exactitud y fiabilidad del clasificador para cada una de las clases, además de los anteriores indicadores globales.

En este caso, y a modo de ejemplo, haremos una lectura de los datos obtenidos tras evaluar los resultados al clasificar la imagen en estudio con el clasificador *Maximum Likelihood*. Si observamos la Tabla 4.6, compuesta por la realidad sobre el terreno a nivel de píxel, por ejemplo, para la clase Arena Húmeda se han clasificado correctamente 901 píxeles, sin embargo 4 han sido clasificados como Arena Seca, 1 como Callaos y 2 como Aluviales. En el resto de clases o píxeles no clasificados (unclassified) el valor ha sido de 0 píxeles.

CLASES	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales	<i>Tamarix Canariensis</i>	<i>Tragacanth Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborescens</i>	TOTAL
Arena Seca	1358	10	0	0	3	15	0	0	0	3	0	0	1389
Arena Húmeda	4	901	0	0	1	2	0	0	0	0	0	0	908
Costra Salina	35	1	38	0	1	0	0	0	0	0	0	0	75
Charca	0	0	0	423	0	0	0	0	0	0	0	0	423
Callaos	0	1	0	0	87	0	0	0	0	0	0	0	88
Aluviales	43	44	0	0	0	97	0	0	0	0	0	0	184
<i>Tamarix Canariensis</i>	0	0	0	0	0	0	206	22	6	0	0	1	235
<i>Tragacanth Moquinii</i>	0	0	0	0	0	0	0	127	0	1	2	6	136
<i>Juncus Acutus</i>	0	0	0	0	0	0	1	0	37	0	0	0	38
<i>Suaeda Mollis</i>	0	0	0	0	0	0	0	0	0	20	4	0	24
<i>Tetraena Fontanesii</i>	17	28	0	0	2	5	0	15	1	14	8	1	91
<i>Launaea Arborescens</i>	4	0	2	0	0	0	5	6	1	2	0	40	60
TOTAL	1461	985	40	423	94	119	212	170	45	40	14	48	3651

Tabla 4.6. Matriz de confusión para el clasificador *Maximum Likelihood*.

Finalmente, se expone en la Tabla 4.7 el porcentaje de aciertos para evaluar los resultados de una manera más cómoda a nivel de clases, empleando la ROI de entrenamiento, la ROI de evaluación y los clasificadores *Maximum Likelihood* y *Support Vector Machine*. En dicha tabla podemos apreciar que con el clasificador *Maximum Likelihood*, la clase con menor porcentaje de aciertos (rojo) ha sido la “*Suaeda mollis*” y por el contrario, la clase con mayor porcentaje de aciertos (verde) ha sido la Charca con un 100%. Para el clasificador *Support Vector Machine* tenemos dos clases, Aluviales y “*Tetraena fontanesii*”, con un 0% de probabilidad acierto. Esto significa que dichas clases no han podido ser clasificadas adecuadamente. Posteriormente comprobaremos si esta situación se repite con el clasificador parametrizado. Por el contrario, las clases con mayor porcentaje de aciertos (verde) han sido la Charca, tal y como sucedió con el clasificador *Maximum Likelihood*, y la clase “*Juncus acutus*”.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Clases	<i>Maximum Likelihood</i>	<i>Support Vector Machine</i>
Arena Seca	92,95%	95%
Arena Húmeda	91,47%	87,11%
Costra Salina	95%	62,50%
Charca	100%	100%
Callaos	92,55%	89,36%
Aluviales	81,51%	0%
<i>Tamarix Canariensis</i>	97,17%	85,38%
<i>Traganum Moquinii</i>	74,71%	80%
<i>Juncus Acutus</i>	82,22%	100%
<i>Suaeda Mollis</i>	50%	47,50%
<i>Tetraena Fontanesii</i>	57,14%	0%
<i>Launaea Arborescens</i>	83,33%	75%

Tabla 4.7. Porcentaje de aciertos, a nivel de clases, tras emplear ambos clasificadores.

Los resultados de la tabla anterior, pueden ser analizados visualmente en la Figura 4.21:

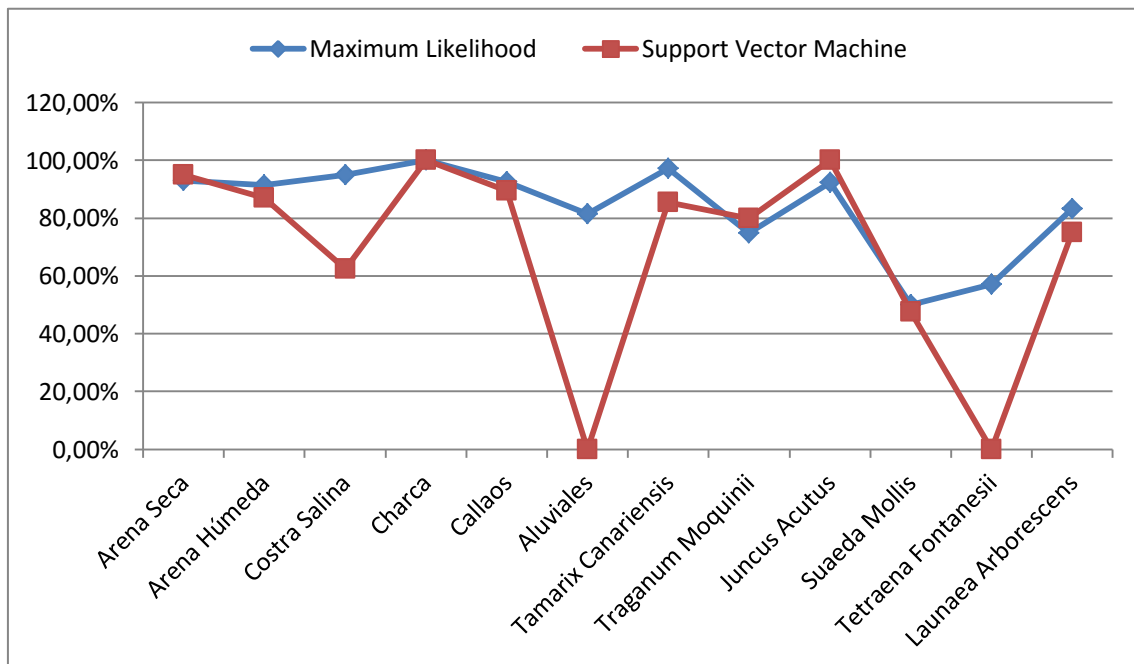


Figura 4.21. Representación de los promedios calculados con los porcentajes de aciertos, a nivel de clase, obtenidos con los clasificadores *Maximum Likelihood* y SVM.

4.4.4. Parametrización y evaluación del SVM

En puntos anteriores de este capítulo hemos desarrollado una serie de clasificaciones empleando los clasificadores *Maximum Likelihood* y *Support Vector Machine*. Una vez realizada la comparativa entre las clasificaciones, vamos a parametrizar el clasificador *Support Vector Machines* para intentar mejorar los resultados obtenidos.

Todas las clasificaciones realizadas hasta el momento han sido con parámetros habituales impuestos por defecto, por lo que nuestro trabajo consistirá en parametrizar y evaluar de forma óptima nuestro clasificador SVM.

Para realizar la parametrización no existe una técnica concreta ni unos valores de parámetros definidos que emplear sino que, todo dependerá del caso que estemos estudiando. Por este motivo, nuestra parametrización tendrá una base teórica que entrelazará las estadísticas y opiniones de diversos autores.

- **Parametrización del SVM**

En primer lugar, debemos tener en cuenta que el hecho de emplear el clasificador SVM implica tomar dos decisiones principales [76]:

- 1) Qué función kernel emplear para mejorar la separabilidad
- 2) Qué metodología de ajuste de parámetros escoger

Dado que en capítulos anteriores hemos detallado las funciones kernel y las metodologías de ajuste, en este caso detallaremos brevemente la información teórica de interés con el fin de centrarnos en la parametrización práctica.

Función Kernel

Basándonos en las conclusiones ofrecidas por diversos autores [97], hemos escogido la **función kernel de base radial** ya que aseguran que ofrece resultados levemente superiores a los de la función kernel polinomial.

En la función kernel de base radial, debemos determinar dos parámetros que además juegan un papel crucial en el rendimiento del SVMs:

- **Parámetro de regularización o costo (C):** Es un valor empleado para ajustar el error de clasificar erróneamente puntos del conjunto de datos de entrenamiento.

- Ancho del kernel (γ): Nos dará una representación de las muestras más plana o abrupta en función de si es un valor pequeño o grande de gamma respectivamente.

Una selección inadecuada de estos parámetros podría causar problemas de sobreajuste por lo que el objetivo es identificar los valores óptimos para ambos parámetros de manera que podamos obtener el mejor *overall accuracy*. Para lograr dicho objetivo, ensayaremos con pares (C, γ) y el primer par con mejor precisión de validación cruzada será seleccionado [98].

Dado que la elección inicial de los parámetros es aleatoria, nos hemos basado en elecciones realizadas por diversos autores, que han tratado con imágenes de teledetección, para hacer nuestra elección [99], [100], [101], [102]. Los valores seleccionados son:

- $C=[2^{-5}, \dots, 2^{21}]$
- $\gamma = [2^{-4}, \dots, 2^3]$

Metodología de ajuste

En capítulos anteriores hemos detallado las metodologías de ajuste de parámetros más empleadas por lo que, en este punto, resumiremos sus características más relevantes para hacer la elección de la metodología a seguir en nuestro proyecto. Anteriormente, hemos mencionado que las metodologías de ajuste de parámetros más empleadas son:

- La búsqueda en red o rejilla (Grid Search).
- La técnica del gradiente descendiente (Gradient Descent Method).

Por un lado, el procedimiento de la técnica Grid Search es el siguiente:

- Se escogen inicialmente de manera aleatoria los valores de los parámetros.
- Se introducen en el clasificador y se evalúa el resultado.
- Se van variando los valores de los parámetros de manera creciente y decreciente y se reevalúan hasta que todos los valores escogidos hayan sido evaluados.
- El par de parámetros (C, γ) que produce el mejor cross-validation accuracy es entonces seleccionado.
- Una vez seleccionemos ese par de parámetros, se hará una búsqueda entre valores más próximos para afinar el resultado.

Por otro lado, tenemos el procedimiento de Gradient Descent Method:

- Es una técnica utilizada cuando tratamos con un gran número de parámetros.
- En primer lugar debemos buscar el espacio de parámetros.
- Se comienza en un punto aleatorio definido por los valores aleatorios de los parámetros.
- Se procede cambiando los valores de los parámetros de acuerdo con la pendiente (derivada) y la dirección de la pendiente de la superficie de error.

Tras varias investigaciones y en concordancia con los estudios realizados por diversos autores [67], [98], [99], [101], [103] y [104], podemos afirmar que la metodología Grid Search es la más utilizada en teledetección por lo que esa será nuestra técnica elegida.

- **Evaluación del SVM para los diferentes parámetros**

A continuación, evaluaremos el clasificador SVM empleando los valores escogidos para los parámetros de la función kernel seleccionada, es decir, evaluaremos SVM empleando la función kernel de base radial con unos valores de $C = [2^{-5}, \dots, 2^{21}]$ y $\gamma = [2^{-4}, \dots, 2^3]$. Las ROIs empleadas para esta evaluación de SVM son la ROI de entrenamiento y de evaluación, definidas al inicio de este capítulo. Dichas ROIs cuentan con píxeles puros y mezclados por lo que son las ROIs más reales que podemos utilizar.

Los resultados que se obtengan, deben ser comparados con los obtenidos tras emplear la ROI de entrenamiento y la ROI de evaluación, así como el clasificador SVM con los parámetros por defecto:

- Función Kernel: RBF (Función Kernel de Base Radial)
- Gamma (γ)=0.125=2⁻³
- Costo (C)= 100

Con estos parámetros, el *overall accuracy* obtenido fué 87.5103% y el kappa coefficient de 0.8301. Nuestro objetivo será superar estos medidores afinando los valores de los parámetros del kernel.

En primer lugar, tomaremos $\gamma = 2^{-3}$ y analizaremos todo el rango de valores de C con saltos de 2². Los resultados obtenidos para el *overall Accuracy* y el coeficiente Kappa (k) se muestran en la Tabla 4.8.

Si analizamos dicha Tabla, podemos ver cómo el valor del *overall accuracy* va en aumento conforme el valor de C se incrementa. Esto sucede hasta que para $C=2^{19}$, alcanzamos el máximo de *overall accuracy* y kappa coefficient con valores de

93.9743% y 0.9192 respectivamente. Este resultado supera, entorno a un 6%, a los resultados obtenidos con los parámetros por defecto. Si nos fijamos en la Tabla 4.8, podemos comprobar que hemos utilizado valores de C que sobrepasan el rango definido inicialmente. Esta decisión ha sido tomada porque el máximo valor de *overall accuracy* se ha obtenido con un valor de C muy cercano al límite del rango por lo que, debemos verificar que no existe un valor superior de C que ofrezca mejores resultados. Efectivamente, podemos observar como ninguno de los nuevos valores de C superan al resultado elegido como óptimo.

Gamma γ	C	Overall Accuracy	Kappa Coefficient	C	Overall Accuracy	Kappa Coefficient
2^{-3}	2^{-5}	83,4566%	0,7711	2^{13}	93,5360%	0,9129
	2^{-3}	84,8261%	0,7914	2^{15}	93,8647%	0,9173
	2^{-1}	86,6886%	0,8179	2^{17}	93,8099%	0,9167
	2	86,9351%	0,8215	2^{19}	93,9743%	0,9192
	2^3	86,1956%	0,8124	2^{21}	93,8647%	0,9176
	2^5	86,5790%	0,8177	2^{23}	93,8647%	0,9176
	2^7	88,7702%	0,8478	2^{25}	93,2895%	0,9098
	2^9	92,8787%	0,9041	2^{27}	93,6182%	0,9144
	2^{11}	92,9882%	0,9055			

Tabla 4.8. Búsqueda gruesa de parámetros. Resultados para kernel RBF y $\gamma = 2^{-3}$.

Los resultados de la Tabla 4.8 se repiten para todo el intervalo de valores de gamma, es decir, para $\gamma=2^{-4}, 2^{-2}, 2^0, 2^1, 2^3$ los valores representados en la Tabla 4.8 se repiten.

Sería precipitado asegurar que el valor de gamma es irrelevante en este estudio por lo que ampliaremos su rango de valores con el fin de asegurarnos de que no nos hemos equivocado en la elección y que efectivamente su valor es irrelevante para la mejora de los resultados.

Previamente, afinaremos los resultados de la Tabla 4.8, es decir, tal y como especifica la teoría de la metodología grid search, debemos realizar una nueva búsqueda del parámetro óptimo pero esta vez con un espaciado más fino entre los valores. Puesto que el máximo lo hemos hallado con $C= 2^{19}$, realizaremos ahora las clasificaciones con

valores de $C = [2^{17}, \dots, 2^{21}]$ y con saltos de $2^{0.25}$. Los resultados para esta nueva búsqueda los hallaremos con el valor que gamma trae por defecto en el clasificador, es decir, para $\gamma = 2^{-3}$. En la Tabla 4.9, se exponen los resultados obtenidos.

Gracias a esta búsqueda fina, el resultado obtenido anteriormente ha sido mejorado aproximadamente un 0.3%. En este caso, el máximo ha sido con $C = 2^{19.5}$.

Gamma γ	C	Overall Accuracy	Kappa Coefficient
2^{-3}	$2^{17.25}$	93,8373%	0,9170
	$2^{17.5}$	93,8647%	0,9174
	$2^{17.75}$	94,0564%	0,9201
	2^{18}	93,8647%	0,9176
	$2^{18.25}$	93,9195%	0,9183
	$2^{18.5}$	93,6730%	0,9150
	$2^{18.75}$	93,8099%	0,9169
	2^{19}	93,9743%	0,9192
	$2^{19.25}$	93,9743%	0,9191
	$2^{19.5}$	94,2482%	0,9229
	$2^{19.75}$	94,1660%	0,9219
	2^{20}	94,1660%	0,9218
	$2^{20.25}$	94,0838%	0,9207
	$2^{20.5}$	94,1934%	0,9221
	$2^{20.75}$	93,9195%	0,9185
	2^{21}	93,8647%	0,9176

Tabla 4.9. Búsqueda fina de parámetros. Resultados para kernel RBF y $\gamma = 2^{-3}$.

Hemos representado en la Figura 4.22 los valores de *overall accuracy* de la Tabla 4.9 para su mejor visualización:

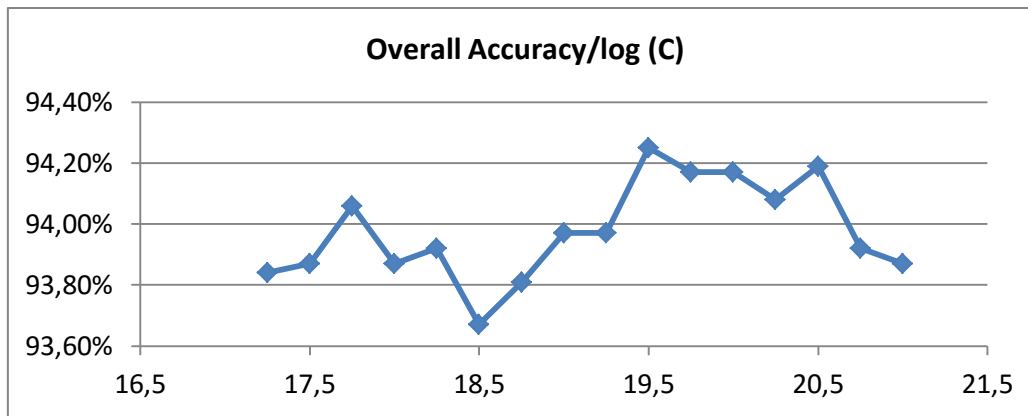


Figura 4.22. Representación gráfica de los resultados de la búsqueda fina.

Resumiendo, hemos tomado como referencia los resultados obtenidos con una parametrización que el clasificador trae por defecto en nuestro programa y a partir de ahí hemos realizado una búsqueda en rejilla. A nivel numérico los progresos han sido los siguientes:

	C	Overall Accuracy	Kappa Coefficient
Parámetros por defecto	100	87,5103%	0,8301
Grid Search Grueso	$2^{19} = 524288$	93,9743%	0,9192
Grid Search Fino	$2^{19,5} = 741455,2$	94,2482%	0,9229

↷ $\approx +6\%$
↷ $\approx +0,3\%$

Tabla 4.10. Resumen de resultados de parametrización.

A nivel de imagen, podemos comprobar las mejoras comparando las Figuras 4.23, 4.24 y 4.25. Si nos fijamos en las Figuras 4.23 y 4.24, observamos como en la primera figura hemos delimitado en el círculo rojo una zona en la que aparece la clase Costra Salina (Blanco). Es una zona en la que debe haber mayoritariamente Arena. Podemos apreciar la gran mejora de la clasificación en esa zona si observamos las Figuras 4.24 y 4.25. Además, podemos apreciar en las zonas delimitadas en azul, como la clase “*Launaea arborescens*”, queda definida de una forma menos compacta conforme mejoramos la clasificación. Este hecho debe darse así, pues dicha especie se encuentra dispersa y no es un área compacta como puede ser la Arena.

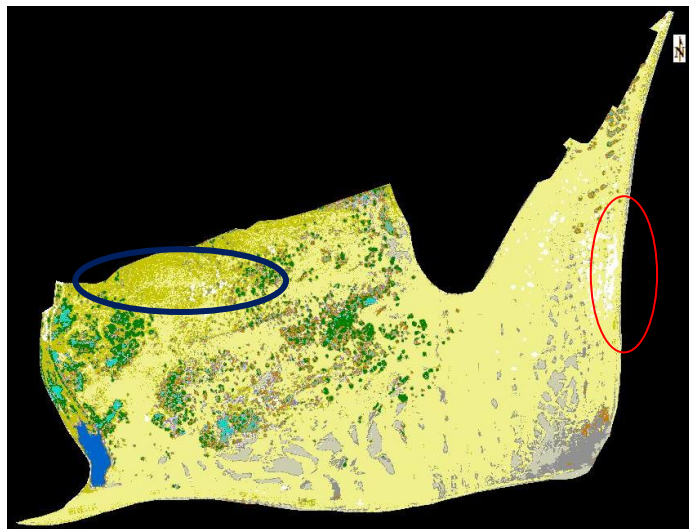


Figura 4.23. Clasificación SVM con parámetros por defecto.

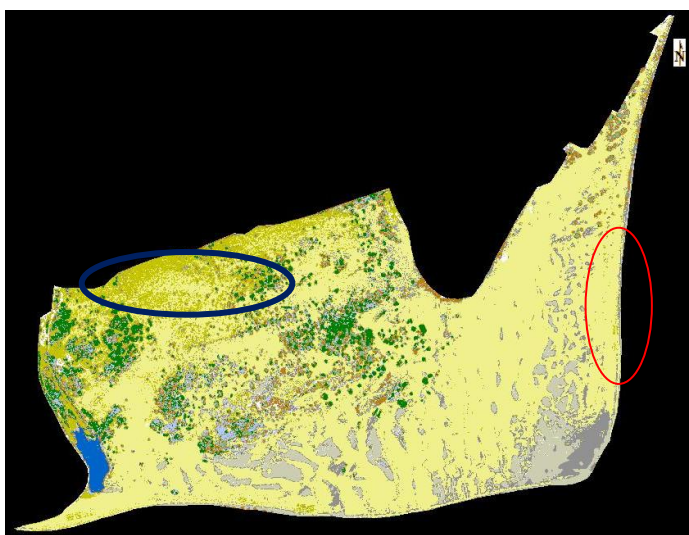


Figura 4.24. Clasificación SVM con Kernel RBF y $C=2^{19}$.

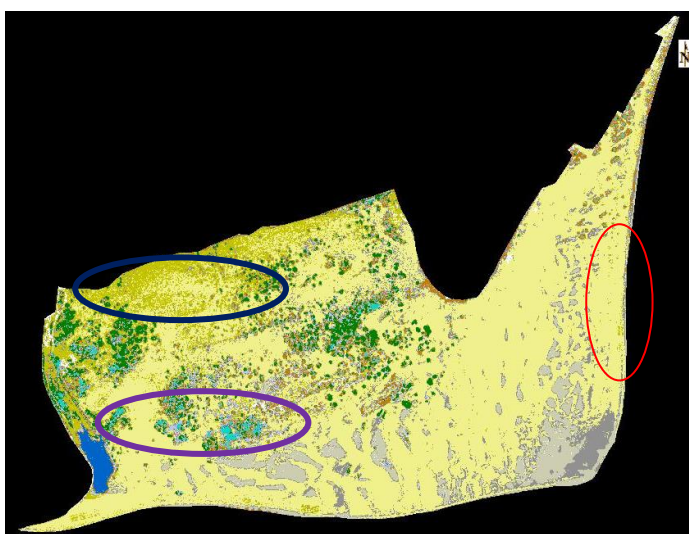


Figura 4.25. Clasificación SVM con Kernel RBF y $C=2^{19.5}$.

Con respecto a la Figura 4.25, si la comparamos con la Figura 4.23 repetiríamos lo indicado en el párrafo anterior ya que es muy similar a la Figura 4.24. Por este motivo, su comparativa con la Figura 4.24 resulta realmente complicada y encontrar una diferencia se convierte en una tarea muy difícil debido a la escasa variación entre sus *overall accuracy* o sus *kappa coefficients*. Una pequeña variación sería la apreciada en la zona violeta, en la cual la clase “*Juncus acutus*” (turquesa) queda mejor definida.

Una vez finalizada la búsqueda del valor óptimo de C y analizados los resultados, vamos a resolver la duda en cuanto al valor de γ .

El valor más pequeño que hemos probado para γ es $2^{-4} = 0.0625$ y el programa ENVI no nos deja seleccionar valores inferiores a 0.010 por lo que las pruebas que vamos a hacer a continuación serán con valores de $\gamma = [2^4, \dots, 2^{22}]$, es decir, superiores a los que ya hemos analizado. Además, para esta prueba fijaremos el valor de C a $2^{19,5}$ ya que ha sido el valor óptimo encontrado para el parámetro de costo. Los resultados obtenidos se muestran a continuación en la Tabla 4.11:

Gamma γ	<i>Overall Accuracy</i>
2^4	94,2482%
2^5	94,2482%
2^6	94,2482%
2^8	94,2482%
2^{10}	94,2482%
2^{12}	94,2482%
2^{14}	94,2482%
2^{16}	94,2482%
2^{18}	94,2482%
2^{20}	94,2482%
2^{22}	94,2482%

Tabla 4.11. Pruebas con el parámetro Gamma (γ).

Podemos observar que siguen sin existir cambios en los valores del *overall accuracy* independientemente de los valores que tome γ . Por este motivo, afirmamos

que en nuestro caso de estudio el valor de γ es irrelevante y daremos por concluida la evaluación de la parametrización del clasificador SVM.

4.4.5. *Linear Spectral Unmixing*

En capítulos anteriores, hemos detallado en qué consiste la técnica de *Linear Spectral Unmixing*. Por este motivo, nos centraremos en analizar los resultados.

Tras aplicar dicha técnica obtendremos un mapa de abundancias de las clases de interés en este Trabajo Fin de Máster. Posteriormente podremos obtener la cartografía correspondiente y evaluarla.

Para obtener el mapa de abundancias, debemos seleccionar nuestra colección de *endmembers* (clases) que, en este caso, será nuestra ROI de entrenamiento generada al comienzo del estudio. Para líneas futuras de trabajo se podrían emplear como *endmembers* las firmas espectrales obtenidas en la campaña de campo del proyecto ARTEMISat

Haciendo uso de la ROI y aplicando la técnica *Linear Spectral Unmixing*, se obtienen los mapas de abundancia. En las Figuras 4.26 (a), (b), (c), (d), (e) y (f) se muestran los mapas a nivel de sustrato: arena seca, arena húmeda, costra salina, charca, callaos y aluviales respectivamente.

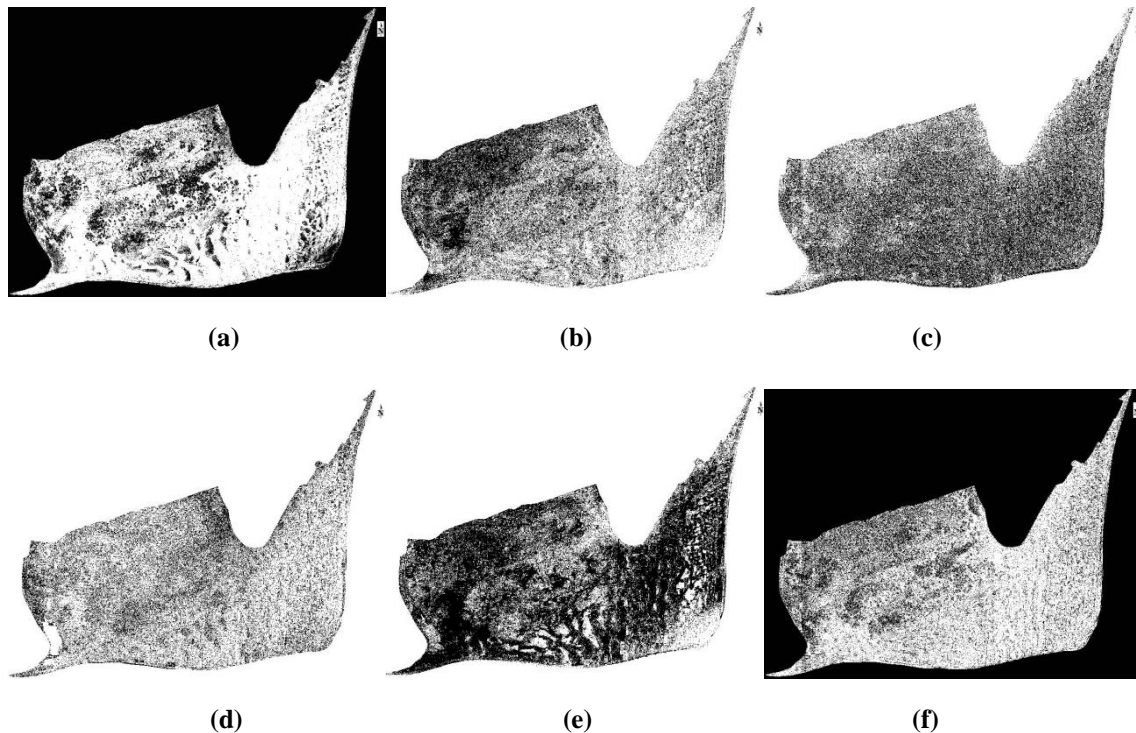


Figura 4.26. Mapas de abundancia para las clases pertenecientes al grupo de sustratos.

Análogamente, en las Figuras 4.27 (a), (b), (c), (d), (e) y (f) se presentan los mapas de abundancia de las clases pertenecientes al grupo de vegetación: “*Tamarix canariensis*”, “*Traganum moquinii*”, “*Juncus acutus*”, “*Suaeda mollis*”, “*Tetraena fontanesii*” y “*Launaea arborescens*” respectivamente.

Si analizamos estos mapas, podemos apreciar que los resultados para la vegetación han sido, en general, peores que los de las clases englobadas en sustrato.

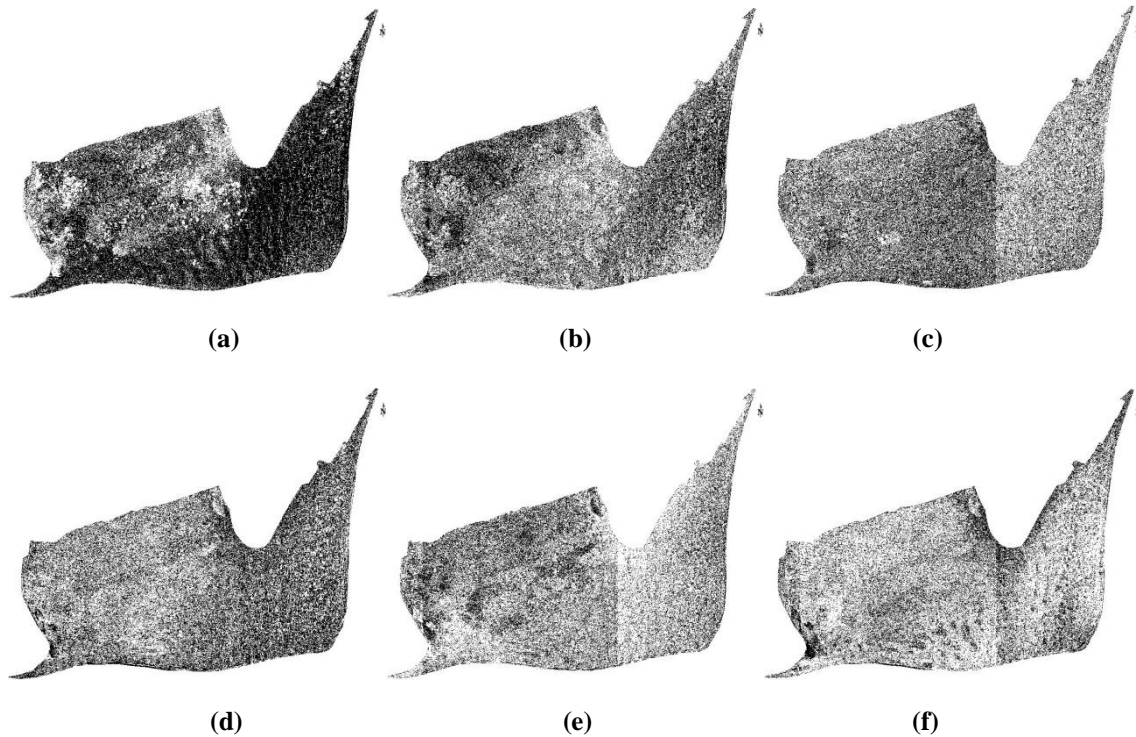


Figura 4.27. Mapas de abundancia para las clases pertenecientes al grupo de vegetación.

Una vez obtenidos los mapas de abundancia, debemos crear un árbol de decisión que nos permita, a partir de esos mapas, obtener la cartografía de la Reserva Natural de las Dunas de Maspalomas. Dicho árbol de decisión se muestra en la Figura 4.28. Para generarlo, se hace una comparativa de cada clase respecto al resto y a nivel de píxel.

Por este motivo, resulta un árbol de $N-1$ nodos principales (padres) siendo N el número de clases. Los colores que determinan a las clases como nodos hijo se corresponden con los asignados en la leyenda de apartados anteriores.

Finalmente, en la Figura 4.29 presentamos la cartografía obtenida para el área de estudio tras ejecutar el árbol de decisión generado. Sin realizar una comparativa, se evidencia el mal resultado.

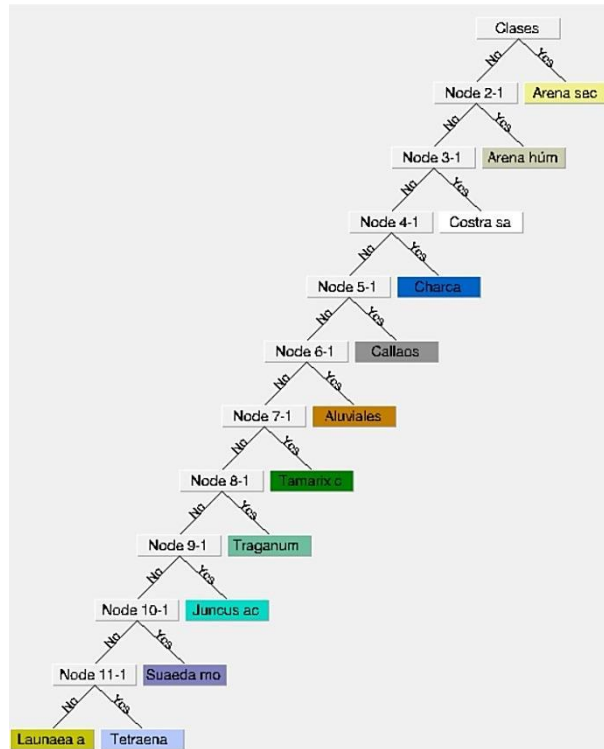


Figura 4.28. Árbol de decisión para generar la cartografía con los mapas de abundancias.

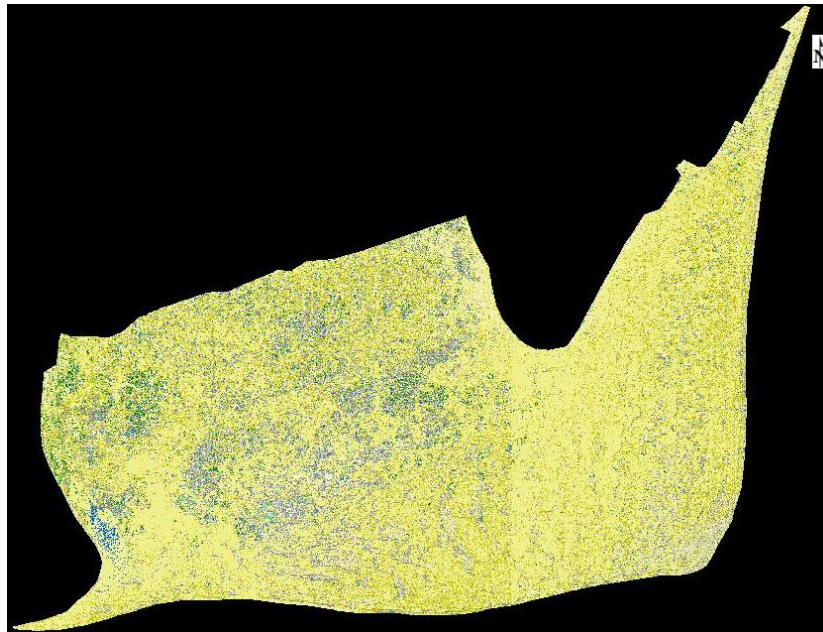


Figura 4.29. Cartografía de la Reserva tras emplear *Linear Spectral Unmixing*.

Visualmente, extraeríamos conclusiones de forma clara y concisa comparando las imágenes de la Figura 4.30 (a), (b), (c) y (d), que se corresponden con el resultado para *Linear Spectral Unmixing*, los resultados de las clasificaciones *Maximum Likelihood* y *Support Vector Machine* optimizado, y el mapa de referencia respectivamente. En casos

como el “*Tamarix canariensis*”, la charca o la arena húmeda, aparecen en las zonas adecuadas de la cartografía pero sin llegar a estar perfectamente definidas.

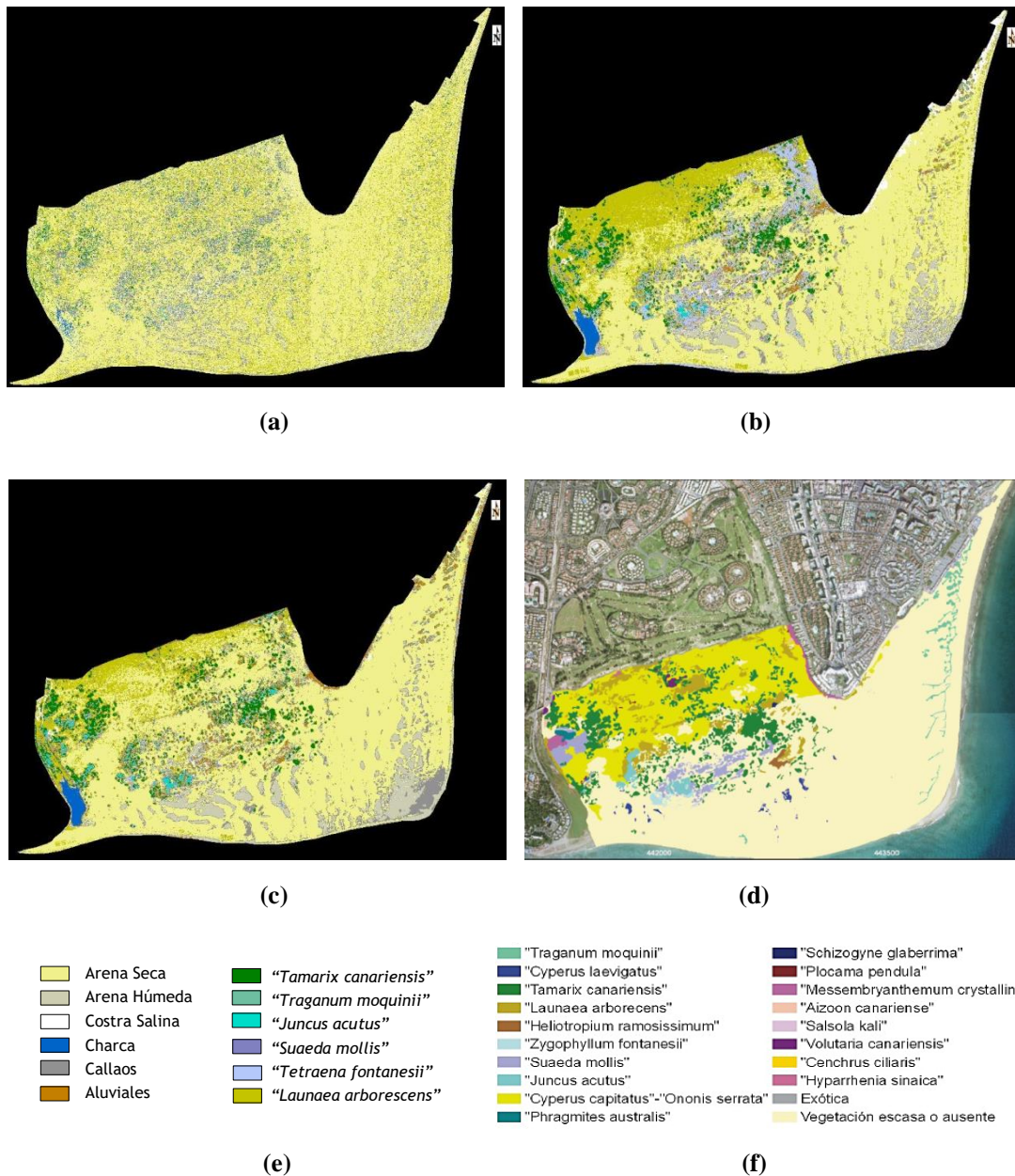


Figura 4.30. Cartografías con: (a) *Linear Spectral Unmixing*, (b) *Maximum Likelihood*, (c) *Support Vector Machine* optimizado, (d) Mapa de referencia, (e) Leyenda de las cartografías generadas y (f) Leyenda del mapa de referencia.

• Evaluación de los resultados de *Linear Spectral Unmixing*

Empleando el mismo procedimiento de apartados anteriores, vamos a evaluar los resultados tras aplicar *Linear Spectral Unmixing* empleando la ROI de evaluación. En la Tabla 4.12 podemos comprobar que el resultado para *overall accuracy* y coeficiente

Kappa son los peores obtenidos hasta el momento y van en concordancia con las conclusiones extraídas visualmente.

	Overall Accuracy	Coefficiente Kappa (K)
Linear Spectral Unmixing	52,8200%	0,2171

Tabla 4.12. Overall accuracy y coeficiente Kappa para *Linear Spectral Unmixing*.

4.4.6. Filtrado *Majority* sobre la imagen original

Llegados a este punto, la última etapa consistirá en filtrar nuestra imagen clasificada con el objetivo de hacer desaparecer los píxeles aislados, si consideramos que son erróneos. Para ello emplearemos el filtrado *Majority* que ha sido explicado brevemente en capítulos anteriores.

Aplicaremos el filtrado *Majority* a la clasificación obtenida para la imagen original, empleando el clasificador *Maximum Likelihood* y el *Support Vector Machine* con sus parámetros óptimos de cara al caso de estudio, es decir, $C=2^{19.25}$, $\text{Gamma}=0.125$ y kernel RBF. Además, filtraremos el resultado obtenido tras emplear la técnica *Linear Spectral Unmixing*.

Para poder observar bien el resultado de aplicar *Majority*, vamos a añadir en la misma figura el resultado con y sin *Majority*. En primer lugar, tenemos la Figura 4.31. En ella tenemos los resultados del clasificador *Maximum Likelihood* sobre la imagen original y enmascarada, con y sin *Majority*. Ese efecto suavizado que se aprecia en la Figura 4.31 (b) es el que se espera tras el filtrado realizado.

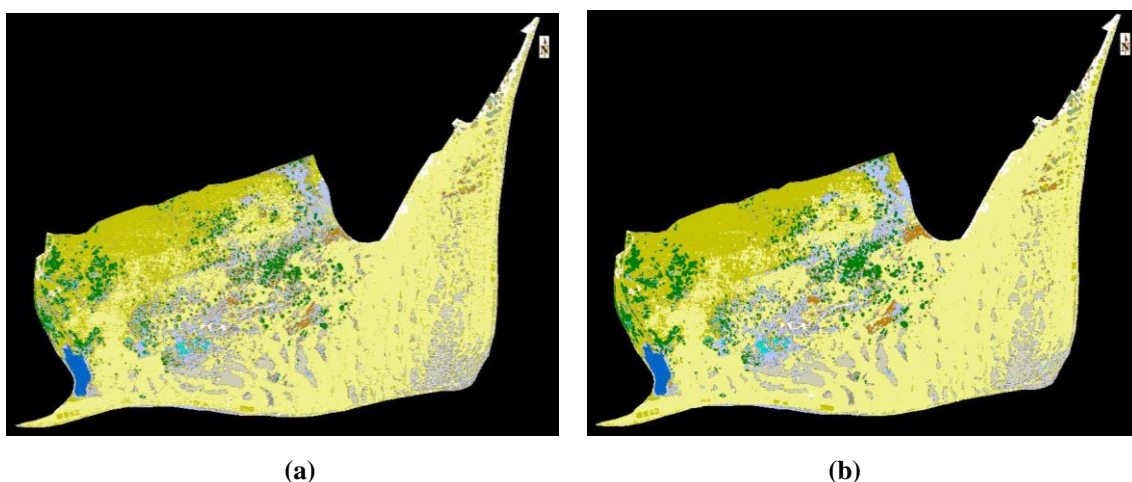


Figura 4.31. Clasificación *Maximum Likelihood* con la Imagen Original enmascarada: (a) sin *Majority* y (b) con *Majority*.

Para obtener un resultado numérico debemos evaluar la clasificación una vez que hemos aplicado el filtro. En la Tabla 4.13 tenemos los resultados de *overall accuracy* y *Kappa* para ambos casos. Si comparamos el *overall accuracy* podemos comprobar que el resultado con el *Majority* es aproximadamente un 3% superior a la clasificación sin filtrado. Por tanto, hemos logrado mejorar la clasificación aplicando un filtrado tras el procesado de la imagen.

	Overall Accuracy	Coefficiente Kappa (K)
Imagen Original sin Majority	91,5366%	0,8885
Imagen Original con Majority	94,3851%	0,9258

Tabla 4.13. Overall Accuracy y Coeficiente Kappa para la imagen original y el clasificador Maximum Likelihood, sin y con Majority.

Tal y como hemos hecho para el clasificador *Maximum Likelihood*, tenemos en la Figura 4.32 el resultado de la clasificación SVM (óptimo) sobre la imagen original enmascarada, con y sin *Majority*. Podemos apreciar como las clases quedan, en general, mejor definidas.

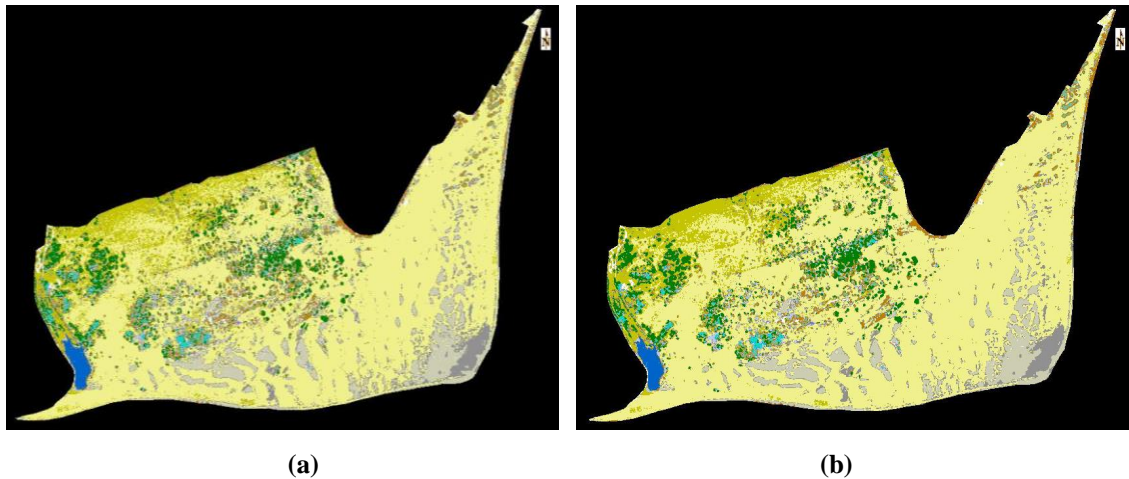


Figura 4.32. Clasificación Support Vector Machine con la Imagen Original enmascarada: (a) sin Majority y (b) con Majority.

Evaluando la imagen filtrada con la ROI de evaluación, podemos apreciar los resultados numéricos comparativos de la Tabla 4.14. La mejora obtenida ha sido superior al 2% por lo que aplicar el filtrado ha sido un procedimiento positivo y mejora el resultado de la clasificación para ambos clasificadores.

	Overall Accuracy	Coefficiente Kappa (K)
Imagen Original sin Majority	94,2482%	0,9229
Imagen Original con Majority	96,7954%	0,9570

Tabla 4.14. *Overall Accuracy* y *Coefficiente Kappa* para la imagen original y el clasificador *Support Vector Machine*, sin y con *Majority*.

Finalmente, aplicamos el filtrado *Majority* a la cartografía obtenida tras aplicar la técnica *Linear Spectral Unmixing* a pesar de los malos resultados obtenidos. En la Figura 4.33 (a) tenemos el resultado original y en la Figura 4.33 (b) el resultado tras aplicar *Majority*. Tal y como ha pasado en los resultados de los clasificadores, se puede apreciar el efecto de suavizado que buscamos con este filtrado pero sin poder determinar que el resultado de esta cartografía es bueno.

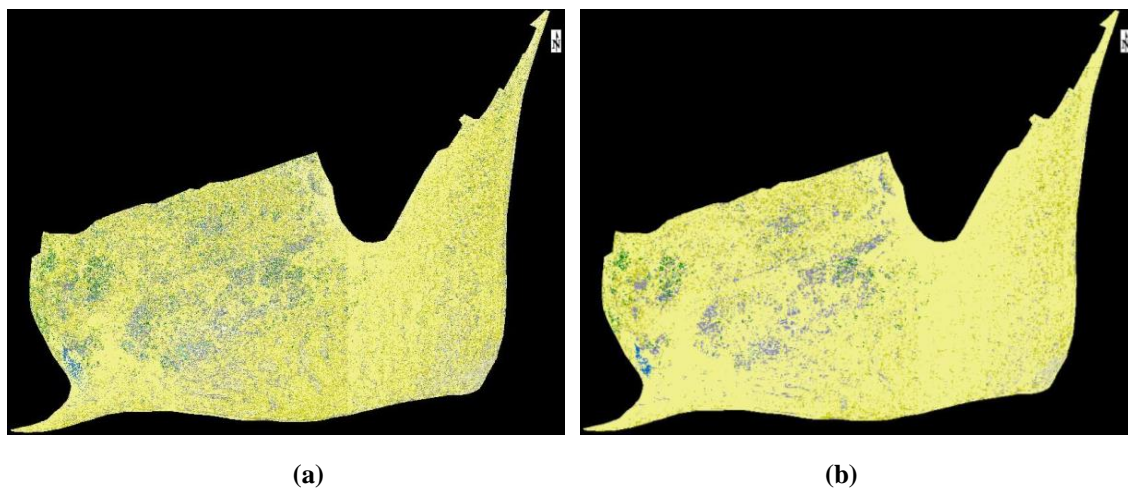


Figura 4.33. *Linear Spectral Unmixing* con la Imagen Original enmascarada: (a) sin *Majority* y (b) con *Majority*.

A pesar del efecto suavizado obtenido, numéricamente se observa en la Tabla 4.15 que el resultado ha empeorado entorno a un 8% respecto al de la imagen original sin *Majority*. Este resultado era el esperado puesto que partíamos de un mapa poco real.

	Overall Accuracy	Coefficiente Kappa (K)
Imagen Original sin Majority	52,8200%	0,2171
Imagen Original con Majority	44,2618%	0,1170

Tabla 4.15. *Overall Accuracy* y *Coefficiente Kappa* para la imagen original y *Linear Spectral Unmixing*, sin y con *Majority*.

4.4.7. Análisis de resultados obtenidos sobre la imagen original

Para poder analizar y comparar los resultados, hemos generado la Tabla 4.16 con los resultados numéricos obtenidos empleando la imagen original. Para extraer conclusiones de una forma más cómoda, se ha añadido la Figura 4.34 el diagrama de barras correspondiente a los *overall accuracy* de la Tabla 4.16. En dicha figura se puede observar que los mejores resultados los ofrece el clasificador *Support Vector Machine* con unos valores ligeramente superiores a los de *Maximum Likelihood*. Además, vemos como el hecho de aplicar el filtrado *Majority* mejora los resultados de ambos clasificadores. Por su parte, la técnica de *Linear Spectral Unmixing* no nos ha ofrecido unos resultados comparables a los de los clasificadores citados anteriormente por lo que su comparativa carece de significado.

	Overall Accuracy	Coefficiente Kappa (K)
Maximum Likelihood	91,5366%	0,8885
SVM óptimo	94,2482%	0,9229
Linear Spectral Unmixing	52,8200%	0,2171
Maximum Likelihood +Majority	94,3851%	0,9258
SVM óptimo + Majority	96,7954%	0,9570
Linear Spectral Unmixing +Majority	44,2618%	0,1170

Tabla 4.16. Overall accuracy y coeficiente Kappa de los resultados sobre la imagen original.

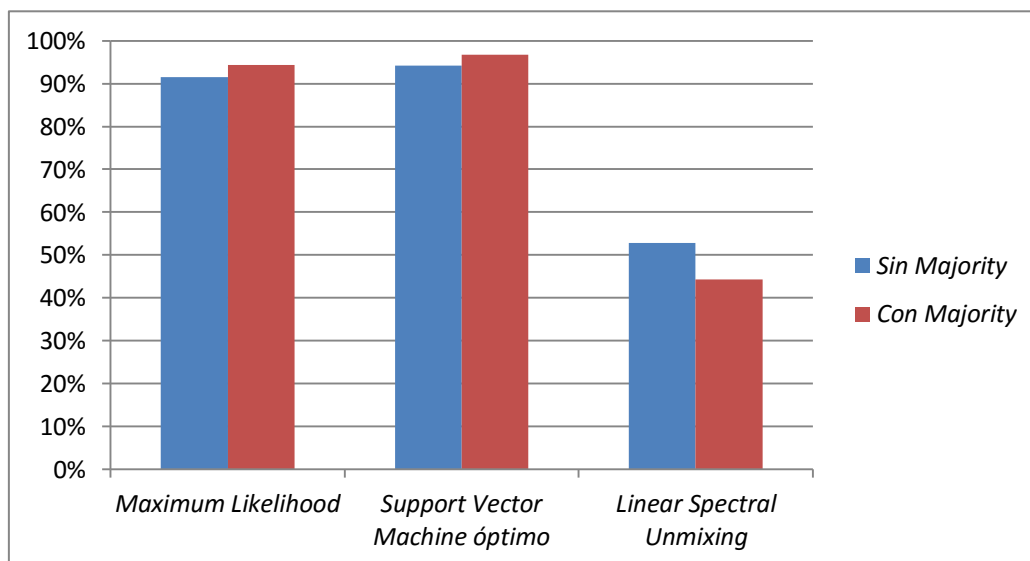


Figura 4.34. Diagrama de barras con los *overall accuracy* de la Tabla 4.16.

Para analizar los resultados a nivel de imagen, tomaremos el mejor resultado obtenido y lo compararemos con el mapa de referencia. Por tanto, vamos a comparar el resultado del clasificador *Support Vector Machine* optimizado y filtrado *Majority* con el mapa de referencia proporcionado por el Doctor Antonio Ignacio Hernández Cordero [93]. Las cartografías y las leyendas correspondientes se exponen en la Figura 4.35.

Puesto que la cartografía de referencia no cuenta con todas nuestras clases del grupo sustrato, hemos añadido en la Figura 4.35 (e) la imagen original de este trabajo para poder comparar y analizar las mismas.

La primera apreciación nos hace concluir comparando los mapas de la Figura 4.35, que las clases mejor discriminadas han sido arena húmeda, arena seca, charca, callos, costra salina, “*Tamarix canariensis*”, “*Juncus acutus*” y la mayor densidad de “*Launaea arborescens*”

En cuanto a los matorrales de “*Suaeda mollis*”, éstos son deficientemente discriminados debido a las propias características de la especie. En muchas ocasiones son clasificados como “*Tamarix canariensis*” porque entre los matorrales de “*Suaeda mollis*” hay ejemplares de esta especie arbórea y por su mayor envergadura ocultan en la imagen al matorral.

Anteriormente hemos dicho que las mayores densidades de “*Launaea arborescens*” quedan perfectamente definidas pero es cierto que los núcleos de las depresiones interdunares de las dunas móviles no son bien diferenciados. Esto puede deberse a que son matorrales de muy poca densidad y no somos capaces de llegar a ese nivel de detalle.

Por su parte, la especie “*Traganum moquinii*” presente en el norte de la Playa del Inglés ha sido discriminada pero únicamente con sus ejemplares de mayor tamaño.

Debemos destacar, que algunas zonas localizadas en las depresiones interdunares de las dunas móviles han sido clasificadas como “*Juncus acutus*” y en realidad son “*Cyperus laevigatus*”. Esta especie no ha sido estudiada en este trabajo pero es normal que se confunda con el “*Juncus acutus*” puesto que son especies con ciertas características parecidas.

Finalmente, la “*Tetraena fontanesii*” no queda bien discriminada porque en algunas zonas, tal y como sucedía con la “*Suaeda mollis*”, se entremezcla con el “*Tamarix canariensis*”.

El análisis detallado de cada una de las especies ha sido posible gracias a la ayuda y experiencia del Doctor Antonio Ignacio Hernández Cordero [93].

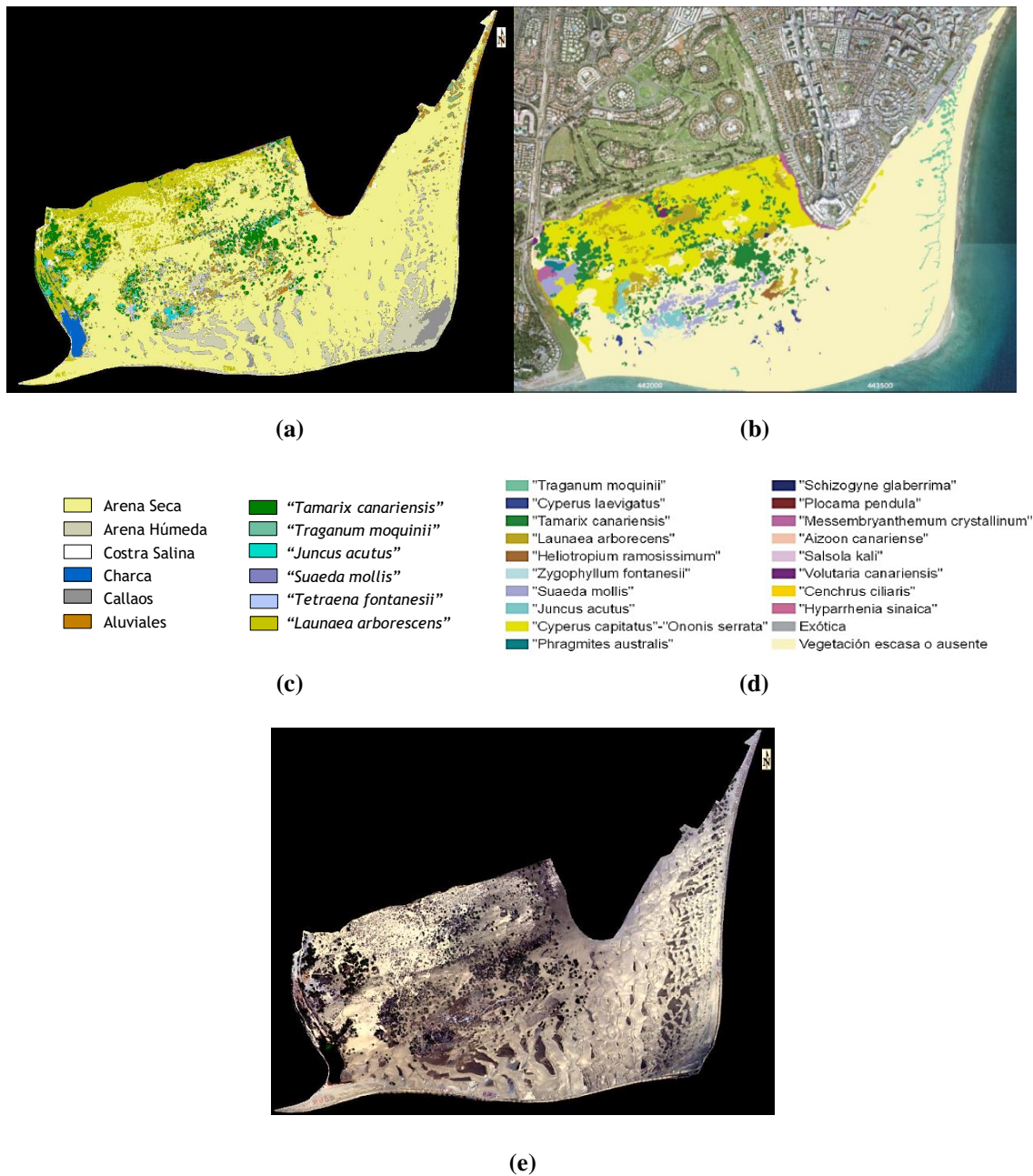


Figura 4.35. Análisis del mejor resultado sobre la imagen original: (a) *Support Vector Machine* optimizado y con filtrado *Majority*, (b) mapa de referencia, (c) Leyenda de la cartografía SVM, (d) Leyenda del mapa de referencia y (e) imagen original.

4.5. Resultados para la imagen fusionada (GS)

Tras realizar el estudio sobre la imagen original multiespectral, nos proponemos mejorar los resultados mediante el uso del Algoritmo de Fusión de Imágenes Gram-Schmidt (GS). En la Figura 4.36 tenemos el esquema completo a seguir para realizar el análisis correctamente.

En capítulos anteriores hemos explicado cómo realizar cada uno de los pasos del esquema por lo que nos centraremos, a continuación, en los resultados obtenidos.

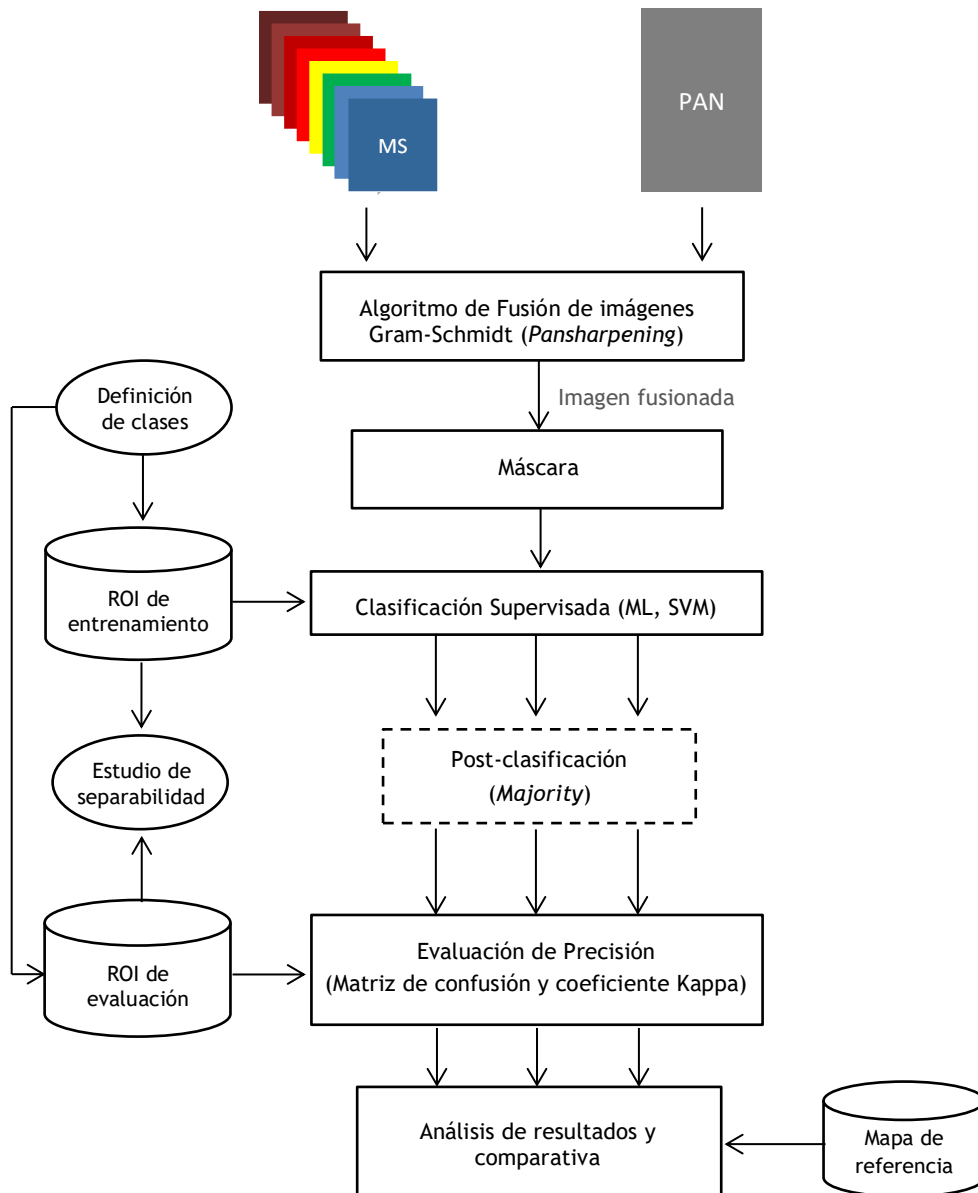


Figura 4.36. Esquema para mejorar la clasificación empleando la fusión de imágenes.

Para que podamos ver la diferencia entre la imagen original y la imagen fusionada hemos añadido la Figura 4.37. En ella podemos observar como el nivel de detalle y claridad de la imagen original es inferior al de la imagen fusionada.

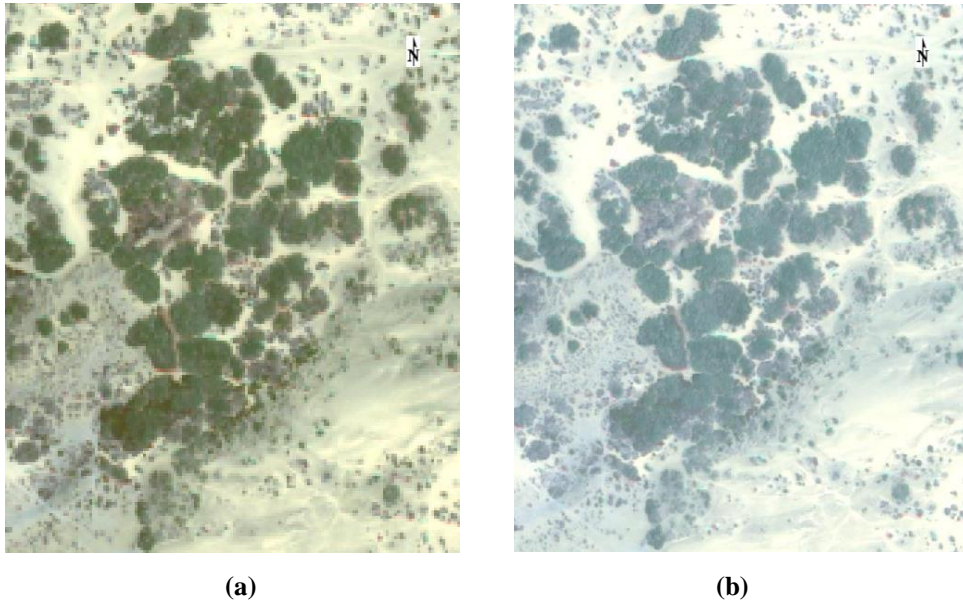


Figura 4.37. Fragmento de la misma zona en (a) la imagen original y (b) imagen fusionada GS.

4.5.1. Regiones de interés (ROIs) y cálculo de la separabilidad

En apartados anteriores de este Trabajo Fin de Máster hemos explicado detalladamente cómo se realizan estas regiones, por lo que, aprovecharemos las mismas ROIs generadas en el apartado anterior y nos centraremos en estudiar la separabilidad de las mismas sobre la imagen fusionada.

4.5.1.1. ROI de entrenamiento para la imagen fusionada

A continuación hemos obtenido, siguiendo el procedimiento explicado en capítulos anteriores, el fichero que nos facilita información relativa a la separabilidad entre clases. Dicha información ha sido dividida en dos partes. La primera parte, formada por los índices de Jeffries-Matusita (JM) y la divergencia transformada que podremos encontrar en el Anexo1. La segunda, se corresponde con los pares de separabilidad obtenidos entre clases. Para una lectura comprensiva de los resultados, hemos plasmado los pares de separabilidad en la Tabla 4.17. Estos valores difieren mínimamente de los obtenidos sobre la imagen original, manteniéndose el peor par de separabilidad entre las clases “*Suaeda mollis*” y “*Tetraena fontanesii*”. Dicho par de separabilidad ha sido remarcado en rojo en la Tabla citada. Por el contrario, hemos remarcado en color verde los pares de

separabilidad con valor óptimo, coincidiendo con lo obtenido para la imagen original. Una vez más, la clase charca obtiene la mejor separabilidad respecto al resto de las clases de estudio en este trabajo.

CLASES	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales	<i>Tamarix Canariensis</i>	<i>Tragranum Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborecens</i>
Arena Seca	1.8626	1.9728	2.0000	1.8074	1.7928	1.9999	2.0000	2.0000	1.9975	1.8952	1.9999
Arena Húmeda		1.9917	2.0000	1.9128	1.9383	1.9998	2.0000	2.0000	1.9972	1.7516	2.0000
Costra Salina			2.0000	1.9899	1.9819	1.9999	2.0000	2.0000	2.0000	1.9875	2.0000
Charca				2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000
Callaos					1.9395	1.9999	2.0000	2.0000	2.0000	1.9621	2.0000
Aluviales						1.9999	2.0000	2.0000	2.0000	1.9406	2.0000
<i>Tamarix Canariensis</i>							1.8322	1.9919	1.9376	1.9554	1.9834
<i>Tragranum Moquinii</i>								1.9997	1.9397	1.9109	1.9445
<i>Juncus Acutus</i>									1.9996	1.9973	1.9988
<i>Suaeda Mollis</i>										1.4208	1.9551
<i>Tetraena Fontanesii</i>											1.9377

Tabla 4.17. Pares de separabilidad para la ROI de entrenamiento (distancia JM) sobre la imagen fusionada.

En la Figura 4.38 hemos representado los pares de separabilidad de la Tabla 4.17. De esta forma hemos comprobado que, con la salvedad del mínimo, los pares de separabilidad superan el valor mínimo recomendado (1.7) para unos buenos resultados.

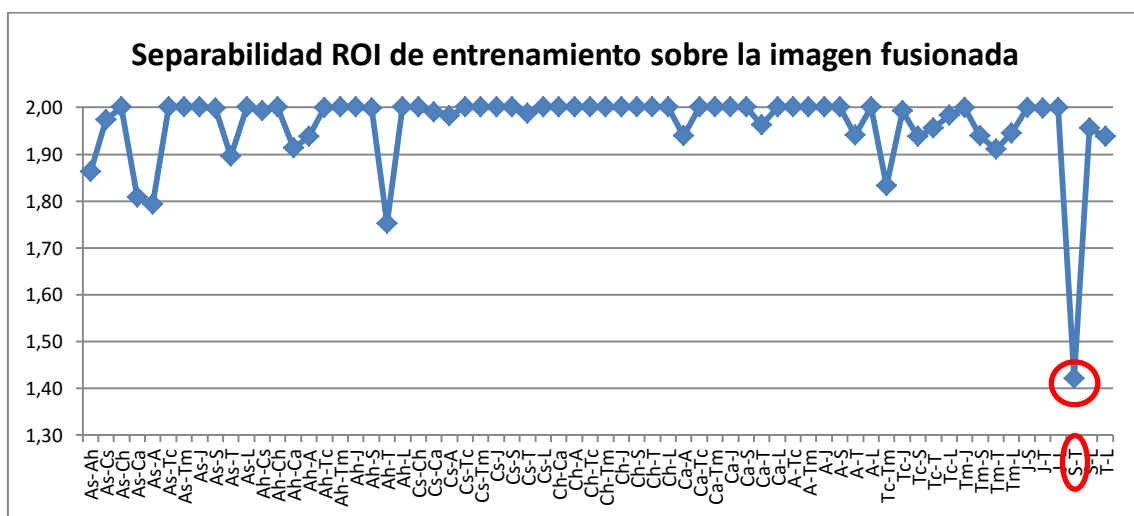


Figura 4.38. Representación de los pares de separabilidad para la ROI de entrenamiento (distancia JM) sobre la imagen fusionada.

4.5.1.2. ROI de evaluación para la imagen fusionada

Llegados a este punto, vamos a definir la región de interés de evaluación con la que mediremos la fiabilidad de los clasificadores tras haberlos entrenado con la región de interés de entrenamiento.

Vamos a analizar los pares de separabilidad de la ROI de evaluación construida en apartados anteriores pero, en este caso, sobre la imagen fusionada.

Los índices de Jeffries-Matusita y de la divergencia transformada para esta ROI de evaluación pueden ser consultados en el Anexo 1. Los pares de separabilidad se pueden consultar en el Anexo 2 y en la Tabla 4.18.

En primer lugar, podemos apreciar que se da el peor par de separabilidad entre las clases “*Suaeda Mollis*” y “*Tetraena Fontanesii*” con un valor de 1.6871. Este par de clases coincide con el peor par obtenido en la ROI de entrenamiento. La diferencia recae en que la distancia JM es algo superior.

CLASES	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales	<i>Tamarix Canariensis</i>	<i>Tragnum Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborecens</i>
Arena Seca	1.9064	1.9636	2.0000	1.8298	1.7984	1.9999	1.9997	2.0000	1.9860	1.9983	1.9999
Arena Húmeda		1.9883	2.0000	1.8456	1.8658	1.9999	1.9999	2.0000	1.9844	1.9886	1.9999
Costra Salina			2.0000	1.9827	1.9759	1.9999	1.9979	2.0000	1.9845	1.9929	1.9991
Charca				2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000	2.0000
Callaos					1.9406	2.0000	1.9998	2.0000	1.9997	1.9979	2.0000
Aluviales						2.0000	1.9989	2.0000	1.9959	1.9826	2.0000
<i>Tamarix Canariensis</i>							1.8174	1.9863	1.9981	1.9781	1.9884
<i>Tragnum Moquinii</i>								1.9985	1.9413	1.8189	1.8750
<i>Juncus Acutus</i>									1.9999	1.9959	1.9968
<i>Suaeda Mollis</i>										1.6871	1.9283
<i>Tetraena Fontanesii</i>											1.9277

Tabla 4.18. Pares de separabilidad para la ROI de evaluación (distancia JM) sobre la imagen fusionada.

Para interpretar cómodamente los pares de separabilidad de la Tabla 4.18, se ha añadido la Figura 4.39. En ella apreciamos cómodamente las conclusiones extraídas anteriormente.

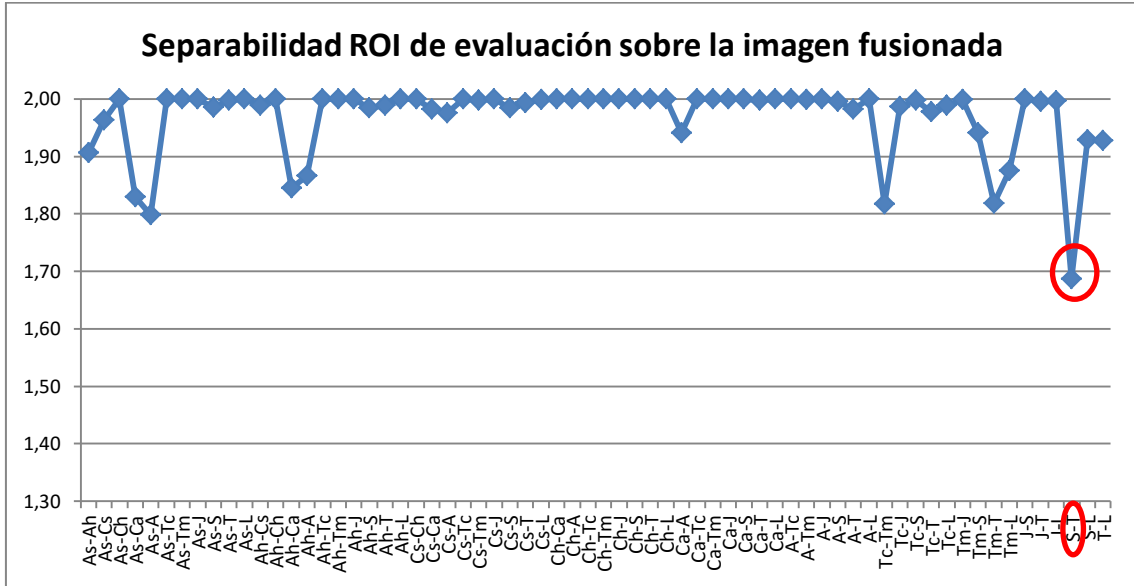


Figura 4.39. Representación de los pares de separabilidad para la ROI de evaluación (distancia JM) sobre la imagen fusionada.

Teniendo en cuenta los resultados obtenidos, hemos representado en la Figura 4.40 los pares de separabilidad, tanto para la ROI de entrenamiento como para la ROI de evaluación. De esta forma, auguramos un buen entrenamiento de nuestros clasificadores y una evaluación robusta de los resultados tras la clasificación.

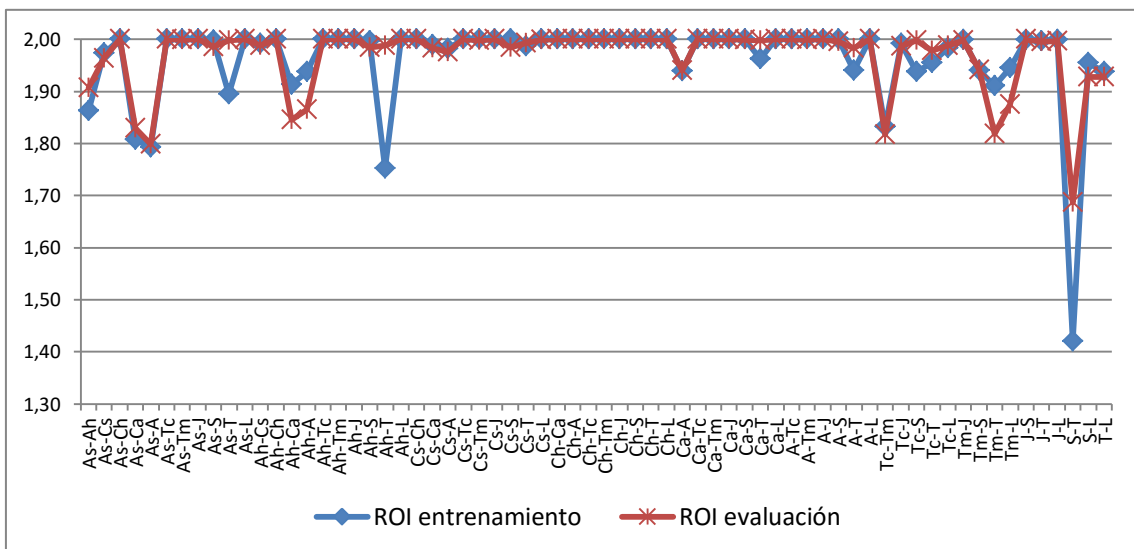


Figura 4.40. Pares de separabilidad de las ROIs de entrenamiento y evaluación (distancia JM) sobre la imagen fusionada.

4.5.2. Generación y aplicación de la Máscara

Teniendo en cuenta que la imagen es de la misma zona, podríamos pensar que la máscara a emplear es la misma que en apartados anteriores. Esto no es así porque las resoluciones de la imagen original y la imagen fusionada son diferentes. Por este motivo, la nueva máscara a emplear es la de la Figura 4.41.



Figura 4.41. Máscara para la imagen fusionada.

4.5.3. Clasificación supervisada y matrices de confusión

En este apartado, distinguiremos por un lado las clasificaciones *Maximum Likelihood* y por otro las realizadas con *Support Vector Machine* de forma que podamos ir comparando los resultados con los de la imagen original sin fusionar.

4.5.3.1. Clasificaciones empleando *Maximum Likelihood* (ML)

En primer lugar hemos realizado la clasificación de la imagen fusionada empleando el clasificador *Maximum Likelihood* y la ROI de entrenamiento adaptada. El resultado obtenido a nivel de imagen para la clasificación se puede apreciar en la Figura 4.42 y además podemos compararlo con el resultado obtenido sobre la imagen original.

Para poder analizar correctamente la clasificación, vamos a evaluarla con la ROI de evaluación adaptada y analizaremos la matriz de confusión resultante.

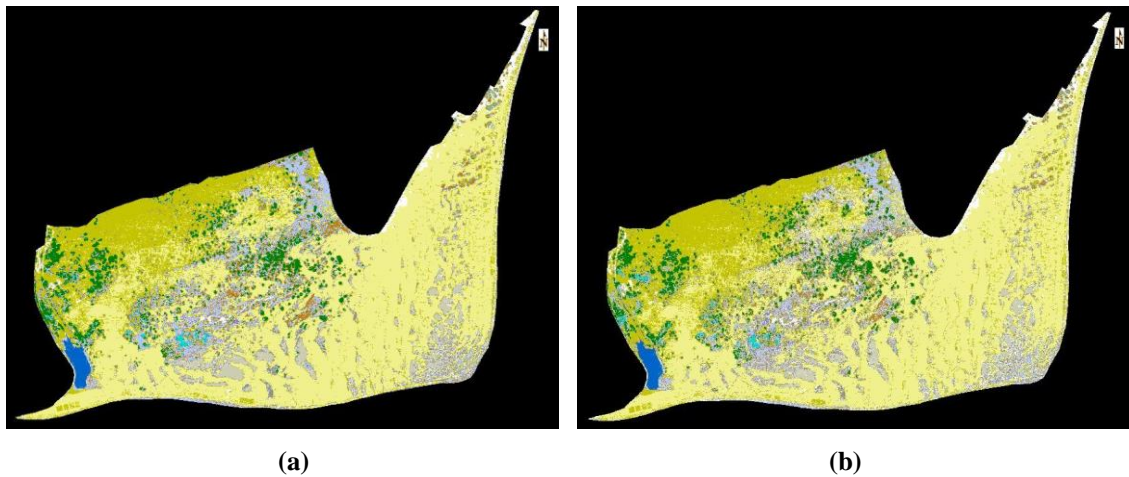


Figura 4.42. *Maximum Likelihood* (a) con la imagen original y (b) con la imagen fusionada.

En la Tabla 4.19 podemos apreciar el resultado de *overall accuracy* y coeficiente Kappa para la imagen fusionada y para la imagen original. Comparando los porcentajes en cada uno de los casos, podemos concluir que el empleo de la imagen fusionada ha hecho que empeoren ligeramente los resultados. En cualquier caso los resultados son muy similares y dependen de las ROIs de entrenamiento y evaluación y para un análisis o representación visual es recomendable el uso de la imagen fusionada pues los contornos de los objetos quedan mejor definidos.

	<i>Overall Accuracy</i>	Coficiente Kappa (K)
Imagen Original	91,5366%	0,8885
Imagen Fusionada	91,0736%	0,8818

Tabla 4.19. *Overall Accuracy* y Coficiente Kappa para las imágenes original y fusionada con ML.

Analizando los valores de la Tabla 4.20 se puede apreciar ese leve descenso de porcentajes en la mayor parte de las clases, con la salvedad de las remarcadas en color verde: Arena Seca, “*Traganum moquinii*”, “*Suaeda mollis*” y “*Launaea arborescens*”.

Hemos añadido la Figura 4.43 para observar de manera más clara la diferencia entre los resultados obtenidos para la imagen original y la fusionada. En ella podemos ver un diagrama de barras que compara los resultados obtenidos para el porcentaje de aciertos a nivel de clase y el *overall accuracy* de la clasificación ML. En la mayor parte de los porcentajes podemos apreciar como la imagen original (azul) supera a la imagen fusionada (rojo).

Clases Sustrato	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales
Imagen original	92,95%	91,47%	95%	100%	92,55%	81,51%
Imagen Fusionada	93,84%	90,52%	92,26%	99,97%	85,97%	61,17%
Clases Vegetación	<i>Tamarix Canariensis</i>	<i>Tragatum Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborescens</i>
Imagen original	97,17%	74,71%	82,22%	50%	57,14%	83,33%
Imagen Fusionada	94,72%	78,46%	81,39%	54,84%	45,54%	89,58%

Tabla 4.20. Porcentaje de acierto a nivel de clase para la imagen original y la fusionada con ML.

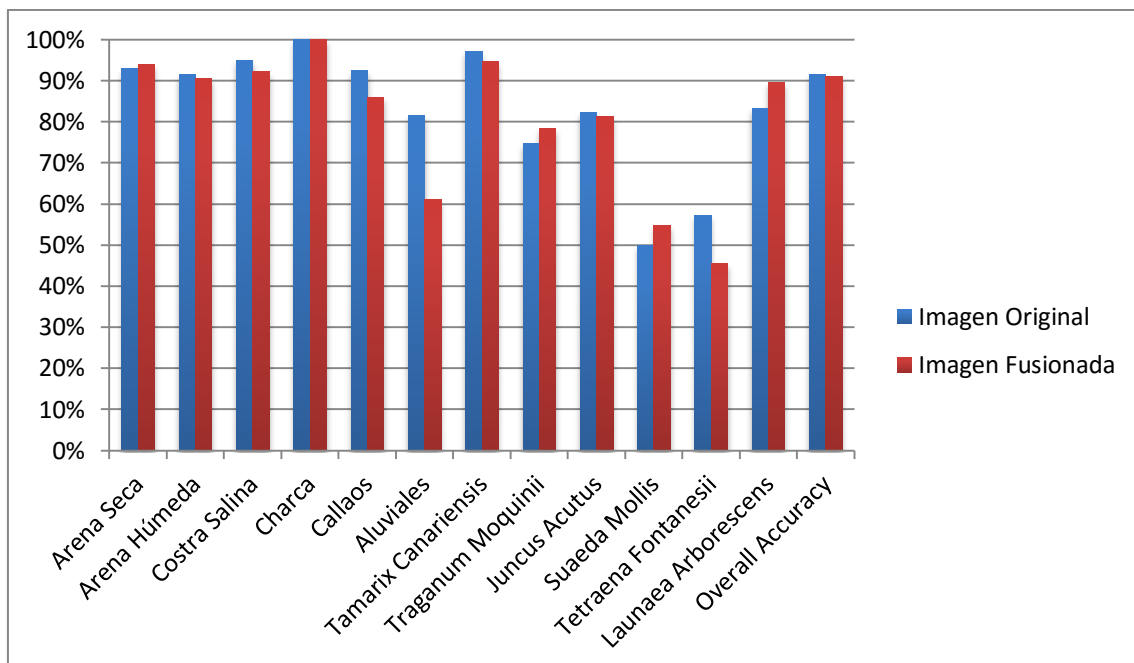


Figura 4.43. Comparativa entre las imágenes original y fusionada para *Maximum Likelihood*.

Con la fusión de imágenes, mejoramos la resolución espacial de nuestra imagen multispectral y eso nos permite ver, de una forma más definida, los detalles de nuestra imagen. Con mayor definición de los elementos de la imagen, deberíamos obtener un resultado superior pero eso no ha sucedido. Esto puede ser debido a que los algoritmos de fusión suelen introducir cierta distorsión espectral (es decir, modifica ligeramente los colores como se aprecian en la Figura 4.37).

4.5.3.2. Clasificación empleando *Support Vector Machine*

A continuación, hemos realizado la clasificación de la imagen fusionada empleando el clasificador *Support Vector Machine* con sus parámetros optimizados y la ROI de

entrenamiento adaptada. El resultado obtenido a nivel de imagen para la clasificación se puede apreciar en la Figura 4.44 y además podemos compararlo con el resultado obtenido sobre la imagen original.

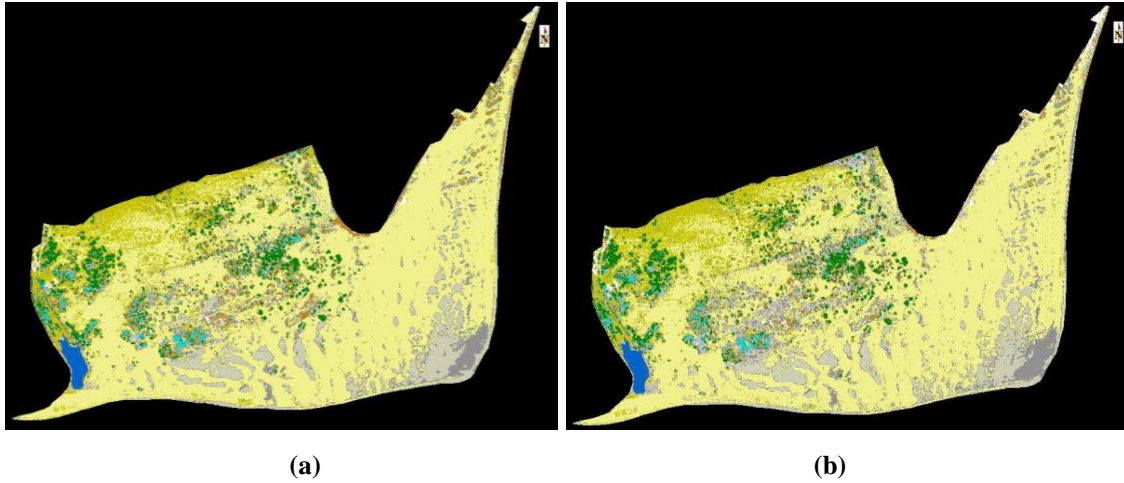


Figura 4.44. Support Vector Machine: (a) con la imagen original y (b) con la imagen fusionada.

Para poder analizar correctamente la clasificación, vamos a evaluarla con la ROI de evaluación adaptada y analizaremos la matriz de confusión resultante.

En la Tabla 4.21 podemos apreciar el resultado de *overall accuracy* y coeficiente Kappa para la imagen fusionada y para la imagen original. Obviamente, en ambos casos hemos empleado el kernel de base radial, con $\text{Gamma} = 0.125$ y $C=2^{19.5}$. Comparando los porcentajes en cada uno de los casos, podemos concluir que el empleo de la imagen fusionada ha hecho que empeoren ligeramente los resultados.

	<i>Overall Accuracy</i>	<i>Coficiente Kappa (K)</i>
Imagen Original	94,2482%	0,9229
Imagen Fusionada	92,0790%	0,8935

Tabla 4.21. Overall Accuracy y Coeficiente Kappa para las imágenes original y fusionada con SVM.

Analizando ahora los porcentajes de acierto a nivel de clase, tenemos la Tabla 4.22. Tal y como hemos hecho con el *overall accuracy* y el coeficiente kappa, añadimos también en la tabla los resultados obtenidos con la imagen original. Si comparamos los valores se puede apreciar ese leve descenso en cada una de las clases salvo en Costra Salina, Callaos, “*Juncus acutus*” y “*Suaeda mollis*”.

Clases Sustrato	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	Aluviales
Imagen original	97,95%	94,72%	70%	100%	93,62%	76,47%
Imagen Fusionada	95,25%	94,49%	77,10%	100%	94,35%	52,63%
Clases Vegetación	<i>Tamarix Canariensis</i>	<i>Traganum Moquinii</i>	<i>Juncus Acutus</i>	<i>Suaeda Mollis</i>	<i>Tetraena Fontanesii</i>	<i>Launaea Arborescens</i>
Imagen original	98,58%	82,94%	86,67%	32,50%	21,43%	87,50%
Imagen Fusionada	94,93%	78,90%	88,47%	44,06%	19,20%	74,61%

Tabla 4.22. Porcentaje de acierto a nivel de clase para la imagen original y la fusionada con SVM.

En la Figura 4.45 se observa de manera más clara la diferencia entre los resultados obtenidos para la imagen original y fusionada. Una vez más, podemos apreciar como la imagen original (azul) supera a la imagen fusionada (rojo) con la salvedad de las clases citadas anteriormente.

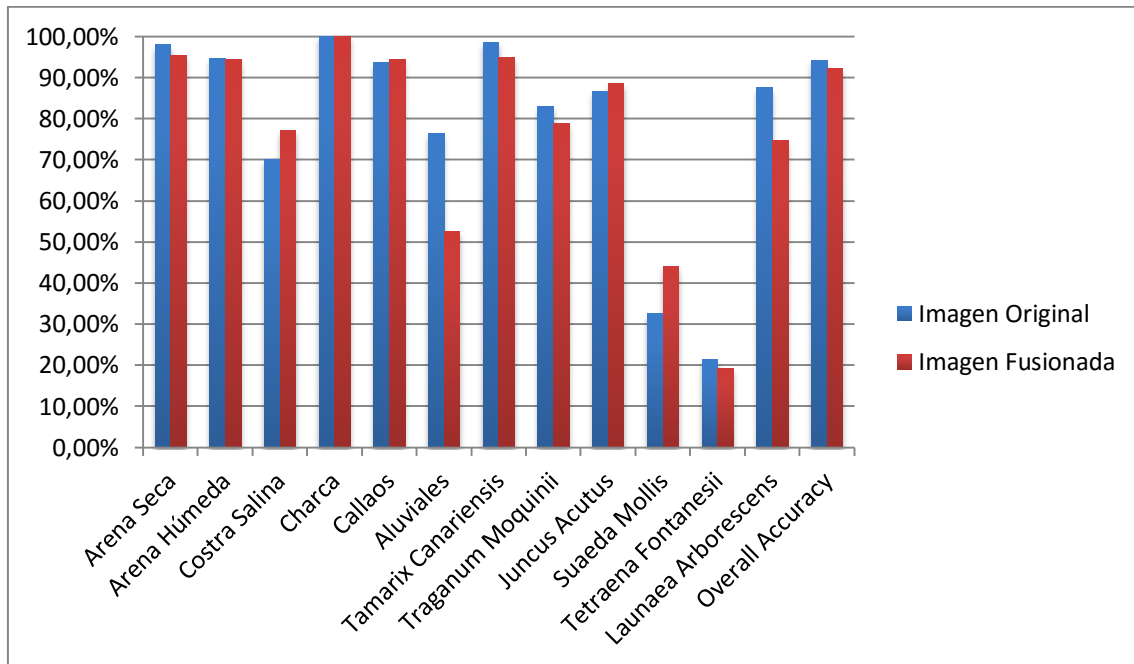


Figura 4.45. Comparativa entre las imágenes original y fusionada para SVM.

Por tanto, y de manera concluyente, afirmamos que en caso de emplear la imagen fusionada para clasificar la Reserva Natural de las Dunas de Maspalomas, los mejores resultados los ha ofrecido el clasificador *Support Vector Machine*. A pesar de esto, realizaremos a continuación un filtrado, tanto al resultado del clasificador *Maximum Likelihood* como al *Support Vector Machine*, para mejorar la representación visual del mapa temático generado.

4.5.4. Filtrado *Majority* sobre la imagen fusionada

Tal y como hemos hecho en el apartado de clasificación, distinguiremos los resultados de cada clasificador por apartados para poder realizar la comparativa respecto a la imagen fusionada sin el filtrado.

4.5.4.1. *Majority* sobre los resultados de *Maximum Likelihood*

A continuación, vamos a analizar los resultados obtenidos con y sin *Majority* pero empleando la imagen fusionada en lugar de la imagen original. Los resultados para el clasificador *Maximum Likelihood* a nivel de imagen se muestran en la Figura 4.46.

En este caso, es más complicado apreciar el trabajo realizado por el filtro *Majority* porque la diferencia visual es inferior que en el caso de la imagen original.

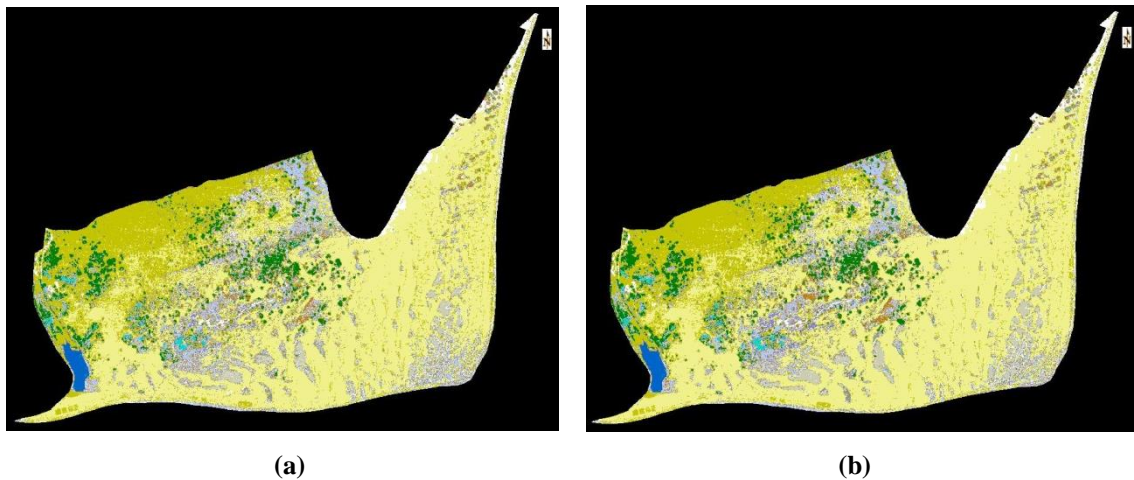


Figura 4.46. Clasificación *Maximum Likelihood* con la Imagen Fusionada enmascarada: (a) sin *Majority* y (b) con *Majority*.

En la Tabla 4.23 podemos comparar los resultados obtenidos en la imagen fusionada para *overall accuracy* y el coeficiente Kappa.

	<i>Overall Accuracy</i>	Coeficiente Kappa (K)
Imagen Fusionada sin <i>Majority</i>	91,0736%	0,8818
Imagen Fusionada con <i>Majority</i>	91,9447%	0,8931

Tabla 4.23. *Overall Accuracy* y Coeficiente Kappa para la imagen fusionada y el clasificador *Maximum Likelihood*, sin y con *Majority*.

La diferencia entre los dos resultados no llega a un 0.9% por lo que es evidente que visualmente es casi imposible encontrar la mejora realizada. A modo de conclusión, podemos afirmar que el hecho de aplicar un filtrado *Majority* mejorará ligeramente los resultados obtenidos para el clasificador *Maximum Likelihood* en la clasificación, sobre la imagen fusionada, de la Reserva Natural de las Dunas de Maspalomas.

4.5.4.2. *Majority* sobre los resultados de *Support Vector Machine*

Analizaremos ahora los resultados obtenidos con y sin *Majority* sobre la imagen fusionada pero con el clasificador *Support Vector Machine*. En la Figura 4.47 podemos apreciar los mapas generados. Los resultados que podemos distinguir parecen menores que los valorados en el caso de la imagen original.

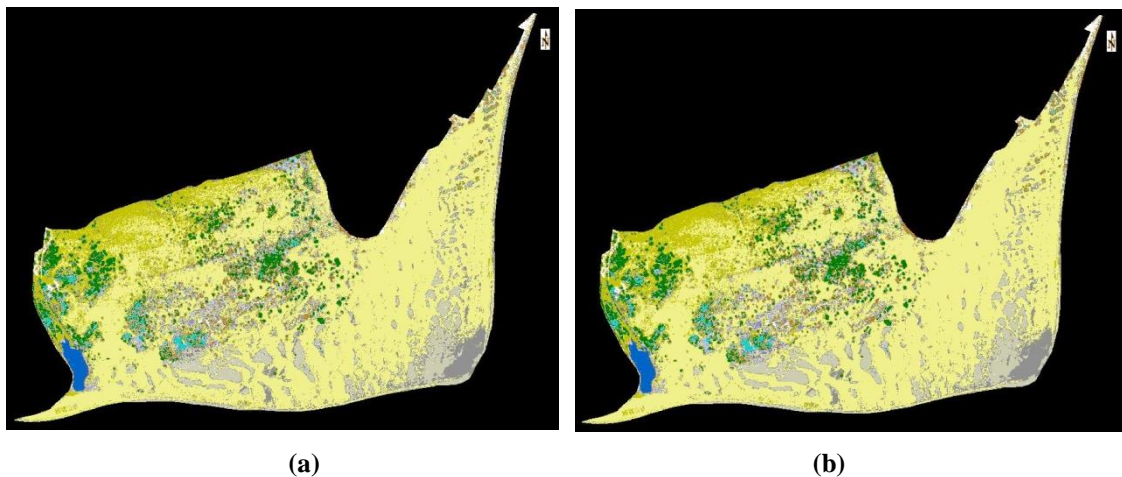


Figura 4.47. Clasificación *Support Vector Machine* con la Imagen Fusionada enmascarada: (a) sin *Majority* y (b) con *Majority*.

Para saber si efectivamente la mejora es inferior, comparamos los resultados de la Tabla 4.24, en la cual se observa una diferencia aproximada de 1,4%. Podemos concluir nuevamente que el empleo del filtrado *Majority* favorece los resultados de la clasificación.

	<i>Overall Accuracy</i>	<i>Coefficiente Kappa (K)</i>
Imagen Fusionada sin <i>Majority</i>	92,0790%	0,8935
Imagen Fusionada con <i>Majority</i>	93,5131%	0,9127

Tabla 4.24. *Overall Accuracy* y *Coefficiente Kappa* para la imagen fusionada y el clasificador *Support Vector Machine*, sin y con *Majority*.

4.5.5. Análisis de resultados sobre la imagen fusionada

Para poder analizar y comparar los resultados, hemos generado la Tabla 4.25 con los resultados numéricos obtenidos empleando la imagen fusionada. Para extraer conclusiones de una forma más cómoda, se ha añadido la Figura 4.48 el diagrama de barras correspondiente a los *overall accuracy* de la Tabla 4.25. En dicha figura se puede observar que los mejores resultados los ofrece el clasificador *Support Vector Machine* con unos valores ligeramente superiores a los de *Maximum Likelihood*. Además, vemos como el hecho de aplicar el filtrado *Majority* mejora los resultados de ambos clasificadores.

	Overall Accuracy	Coefficiente Kappa (K)
Maximum Likelihood	91,0736%	0,8818
SVM óptimo	92,0790%	0,8935
Maximum Likelihood +Majority	91,9447%	0,8931
SVM óptimo + Majority	93,5131%	0,9127

Tabla 4.25. Overall accuracy y coeficiente Kappa de los resultados sobre la imagen fusionada.

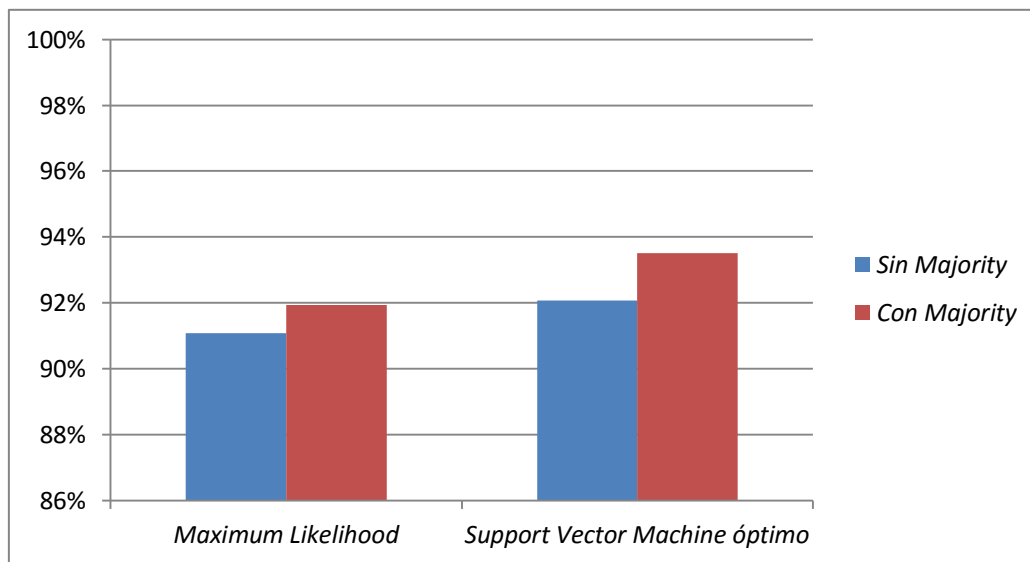


Figura 4.48. Diagrama de barras con los *overall accuracy* de la Tabla 4.25.

Dado que las diferencias a nivel de porcentaje empleando la imagen fusionada son mínimas, no haremos una comparativa a nivel de imagen puesto que sería muy complicado apreciar las diferencias.

4.6. Comparativa de resultados obtenidos

Para poder apreciar de manera apropiada los resultados obtenidos tras clasificar la Reserva Natural de las Dunas de Maspalomas, es conveniente representar de manera conjunta cada uno de los estudios desarrollados.

Hemos representado los resultados obtenidos sobre la imagen original y la imagen fusionada, tanto con el clasificador *Maximum Likelihood* como con el *Support Vector Machine* de parámetros optimizados. Además, se han añadido los resultados obtenidos con el filtrado *Majority*. En la Figura 4.49 se ha plasmado un diagrama de barras con los valores de *overall accuracy* para cada uno de los casos.

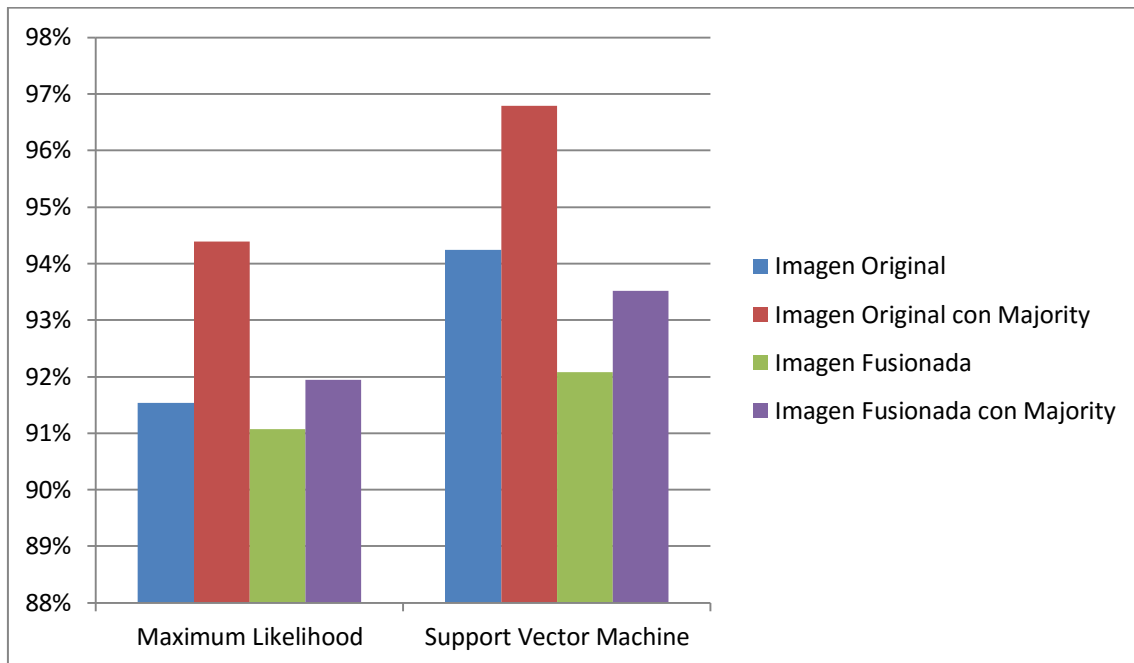


Figura 4.49. Comparativa entre los clasificadores *Maximum Likelihood* y *Support Vector Machine* para la clasificación de la Reserva Natural de las Dunas de Maspalomas.

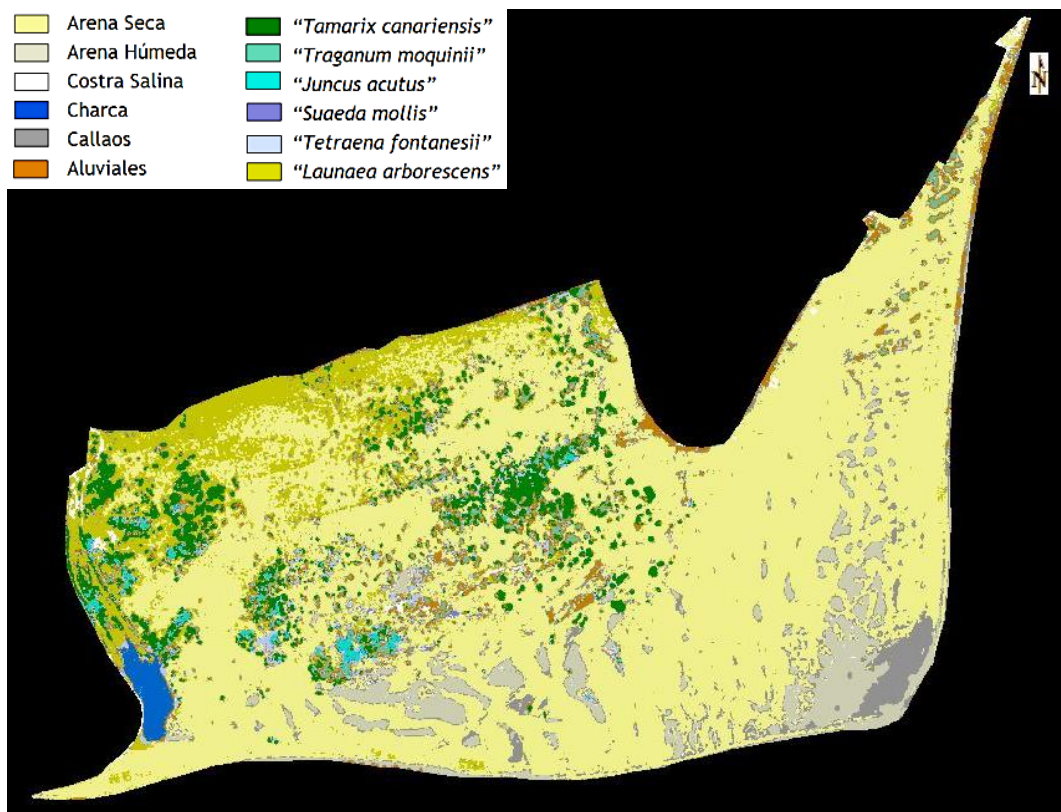
En primer lugar, podemos comprobar en ambos clasificadores como el hecho de aplicar el filtrado *Majority* mejora en mayor o menor medida todos los resultados, tanto para la imagen original como para la imagen fusionada. Dicha apreciación se obtiene comparando entre sí, para cada clasificador, las barras azul y morada por un lado, y las barras verde y violeta por otro.

Además, se puede observar que el clasificador *Support Vector Machine* supera en todos los casos al clasificador *Maximum Likelihood* con independencia de la utilización de la imagen original o la imagen fusionada.

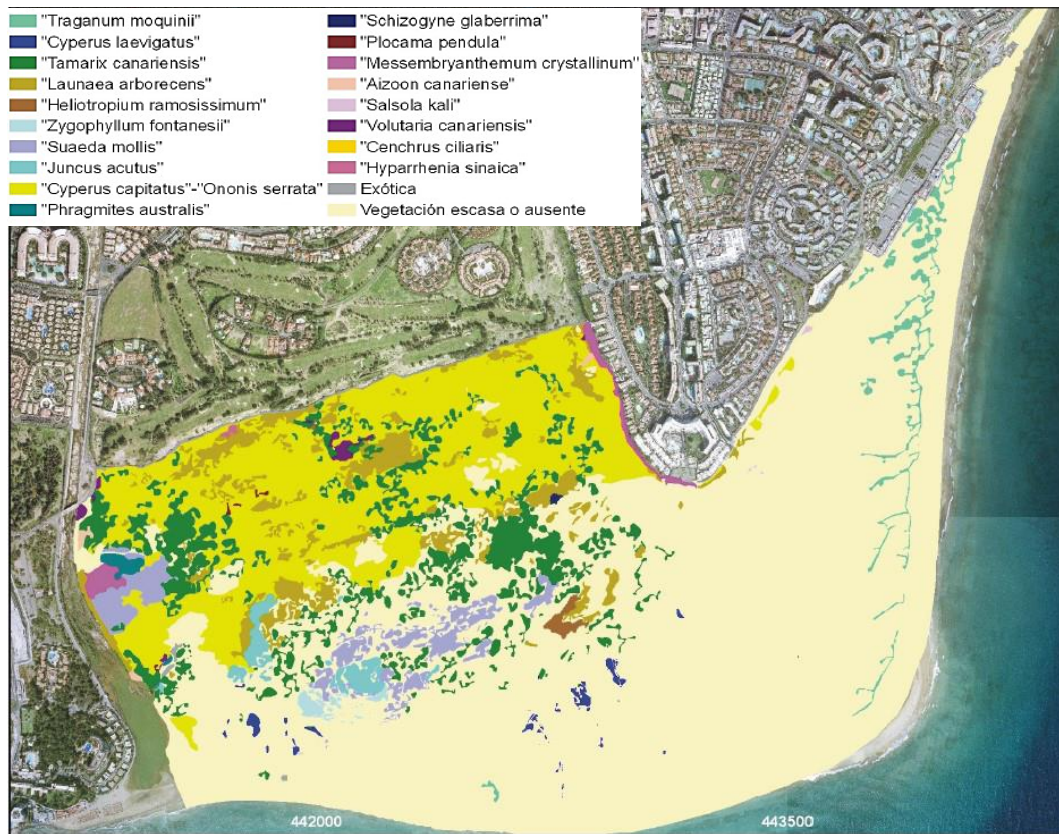
Con toda esta información, determinamos que la metodología óptima y robusta para clasificar el sustrato y la vegetación de la Reserva Natural de las Dunas de Maspalomas es:

1. Emplear la imagen original multispectral.
2. Corregirla radiométrica y atmosféricamente.
3. Enmascararla para trabajar únicamente en la zona de interés.
4. Generar las ROIs de entrenamiento y evaluación
5. Clasificar la imagen empleando *Support Vector Machine* optimizado.
6. Aplicar el filtrado *Majority*

El resultado a nivel de imagen de aplicar dicha metodología es el que se observa en la Figura 4.50 (a). De forma comparativa y avalando nuestro resultado, añadimos en la Figura 4.50 (b) el mapa de referencia empleado a lo largo de este Trabajo Fin de Máster.



(a)



(b)

Figura 4.50. Cartografía: (a) óptima de este Trabajo Fin de Máster y (b) de referencia.

- **Tiempos de cómputo**

Una vez demostrado que el clasificador SVM (*Support Vector Machine*) es el que mejor resultados ofrece, debemos centrarnos en estudiar los tiempos de cómputo empleados para obtener las clasificaciones. A continuación, analizaremos la Tabla 4.26 que está compuesta por los tiempos de cómputo calculados para los clasificadores empleados y ambas imágenes, original y fusionada, así como la ROI de entrenamiento.

Clasificador	Tiempo de cómputo (Imagen Original)	Tiempo de cómputo (Imagen Fusionada)
<i>Maximum Likelihood</i>	45sg	25min
<i>Support Vector Machine</i>	38min 27sg	33h 15min

Tabla 4.26. Tiempos de cómputo para los clasificadores empleados.

Si comparamos los tiempos de cómputo de los clasificadores, tenemos una diferencia de 37min 42sg para la imagen original, y una diferencia de 32h 50min para la imagen fusionada. Por tanto, es evidente que el hecho de emplear la imagen fusionada no sólo nos ofrece resultados ligeramente inferiores sino que además conlleva un

tiempo de cómputo que supera en aproximadamente 1 día y medio al de la imagen original.

A pesar de esto, en el estudio que venimos desarrollando a lo largo de este proyecto, no nos veremos afectados por los tiempos de cómputo ya que lo que buscamos es obtener el mejor resultado de clasificación y no una clasificación de menor duración además de menos eficiente. Esta búsqueda es debida a que la aplicación final de nuestro proyecto no requiere una clasificación en tiempo real. Por tanto, tras analizar los tiempos de cómputo y concretar el objetivo de nuestro estudio, nos decantamos por el clasificador SVM como la mejor opción de cara a lograr los mejores resultados.

4.7. Resumen

En este capítulo se han desarrollado, de manera práctica y empleando el software ENVI, cada uno de los capítulos teóricos explicados anteriormente. Se ha expuesto la imagen WorldView 2 empleada a lo largo trabajo, las clases de especial interés que se pretendían discriminar en la clasificación de la Reserva Natural de las Dunas de Maspalomas, las medidas de campo que nos servirán para definir dichas clases y finalmente las cartografías existentes del área de estudio. A continuación, se ha detallado la metodología seguida en el desarrollo de este capítulo. Además, debíamos conocer el software en el que se ha desarrollado este Trabajo Fin de Máster: ENVI 5.0. Dicho software nos ha permitido desarrollar dos vertientes de trabajo, es decir, con la imagen original multiespectral y con la imagen fusionada. Se expusieron los resultados a nivel de etapas, se analizaron y evaluaron con los procedimientos correspondientes y finalmente se compararon con los mapas de referencia. Por último, se realizó una comparativa de los resultados de ambas vertientes y se pudo concluir cual había sido la metodología óptima para obtener la mejor cartografía de la Reserva Natural de las Dunas de Maspalomas a nivel de sustrato y vegetación.

Capítulo 5: Revisión de objetivos y resultados

En este capítulo se revisan los objetivos iniciales y se exponen las conclusiones alcanzadas tras la realización del Trabajo Fin de Máster, así como algunas de las dificultades más relevantes. También se proponen las posibles líneas futuras de trabajo.

5.1. Revisión de objetivos

La línea de trabajo principal de este Trabajo Fin de Máster ha estado orientada al desarrollo de una metodología para realizar la clasificación, a nivel de sustrato y vegetación, de la Reserva Natural de las Dunas de Maspalomas (Gran Canaria) con la ayuda de la imagen de alta resolución obtenida por medio del satélite WorldView 2. Para ello se ha realizado un estudio exhaustivo que ha permitido seleccionar la metodología completa de clasificación de imágenes multiespectrales en datos procedentes de sensores de muy alta resolución espacial.

Para la consecución de este objetivo global ha sido necesario abordar con éxito las siguientes actuaciones:

- ✓ Se ha realizado un estudio del satélite WorldView 2, sus características técnicas y productos.
- ✓ Se han analizado las diferentes técnicas de pre-procesado habituales en imágenes de alta resolución para ser posteriormente aplicadas a las bandas del satélite WorldView 2, en especial la corrección radiométrica y la atmosférica basada en el modelo 6S.
- ✓ Se han estudiado las técnicas de fusión a nivel de pixel para mejorar la resolución espacial de las bandas multiespectrales a partir de la información de la imagen pancromática. De entre las técnicas existentes se aplicó el algoritmo de Gram-Schmidt.
- ✓ Se ha estudiado la metodología completa de clasificación supervisada de imágenes. Analizando en detalle cada una de sus fases: definición de regiones de interés y medida de separabilidad, fase de entrenamiento, fase de asignación a

partir de algoritmos clásicos y avanzados (como el *Support Vector Machine* y *Maximum Likelihood*), fase de postclasificación utilizando el filtro *Majority* y fase de evaluación de la precisión del mapa temático obtenido.

- ✓ En particular, y dado que SVM ofrecía altas prestaciones, se abordó su estudio detallado para seleccionar adecuadamente el kernel y las estrategias de parametrización adecuadas para la extracción de los parámetros óptimos.
- ✓ Además, hemos aplicado técnicas de procesado de imágenes como LSU (*Linear Spectral Unmixing*) con el objetivo de mejorar los resultados ofrecidos por lo clasificadores citados anteriormente.
- ✓ Para la realización del proyecto se ha utilizado una herramienta software avanzada de procesado de imágenes, por este motivo ha sido necesario realizar un estudio de las funcionalidades de la herramienta ENVI 5.0 para poder utilizarla para la generación de la cartografía de vegetación y sustrato a partir de la aplicación de todas las etapas de procesado necesarias en la metodología desarrollada.
- ✓ Aplicando el software ENVI a los datos WorldView 2 se ha hecho un análisis exhaustivo de la robustez de los diferentes algoritmos de clasificación. También se ha verificado la mejora o no existente al incluir bandas fusionadas.

En definitiva, tras este Trabajo Fin de Máster se ha logrado identificar la metodología y los algoritmos más adecuados para el procesado de los datos de alta resolución que permitan la obtención de mapas temáticos de la máxima precisión posible.

5.2. Revisión de resultados

Tras analizar los resultados obtenidos, nos vemos en disposición de listar las principales conclusiones extraídas:

- ✓ ENVI 5.0 es una herramienta muy potente para la clasificación de imágenes de alta resolución.
- ✓ La utilización de una máscara para la clasificación de la Reserva Natural de las Dunas de Maspalomas nos permite aislar todas las zonas de la imagen que no pertenezcan a la misma, lo cual facilita el proceso de clasificación y reduce tiempos de cómputo.

- ✓ Las clases que presentan una mayor similitud espectral (es decir, una peor separabilidad) presentan una mayor dificultad de discriminación.
- ✓ Con las clasificaciones realizadas hemos podido corroborar que el clasificador *Support Vector Machine* es muy robusto y ofrece unos resultados prácticamente inmejorables.
- ✓ La parametrización del clasificador SVM ha permitido mejorar la exactitud global en más de un 6%.
- ✓ A pesar de los excelentes resultados del clasificador SVM y de su robustez, presenta los peores tiempos de cómputo.
- ✓ Emplear la técnica LSU (*Linear Spectral Unmixing*) nos ha ofrecido unos malos resultados empleando nuestra región de interés como *endmembers* por lo que en este caso, descartamos sus resultados.
- ✓ El empleo de técnicas de fusión de imágenes en este caso de estudio ofrece mejoras en la calidad visual pero empeora ligeramente los resultados y además conlleva un trabajo tedioso, complejo y de mayor tiempo de cómputo.
- ✓ La última etapa de nuestro proyecto, el filtrado *Majority*, realiza un suavizado de nuestra imagen clasificada y elimina el problema de los píxeles aislados, mejorando así el resultado final y por tanto siendo productiva su utilización.

En definitiva, la mejor metodología para la clasificación de la Reserva Natural de las Dunas de Maspalomas ha sido utilizando el clasificador SVM con su parametrización óptima sobre la imagen original y con el filtrado *Majority*.

5.3. Líneas futuras de trabajo

Siguiendo en la línea de nuestras conclusiones, los peores resultados han sido los obtenidos con la técnica LSU (*Linear Spectral Unmixing*). Por este motivo, una buena línea de trabajo futura sería realizar un estudio en profundidad de dicha técnica. Para ese estudio se podrían aprovechar como *endmembers* las firmas espectrales recogidas en la campaña de campo realizada dentro del proyecto ARTeMISat.

Dada la importancia de preservar la Reserva Natural de Las Dunas de Maspalomas, se podría realizar una clasificación temporal para comprobar el estado de las dunas, de su vegetación y de la charca con el paso del tiempo. Para ello, se podrían emplear varias imágenes WorldView 2 de estaciones y épocas temporales diferentes.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Debido al coste de las imágenes WorldView 2, únicamente se ha trabajado con la imagen de Junio de 2015. Como línea futura de trabajo se podría plantear un proyecto en el cual se obtuviesen las imágenes de alta resolución por medio de un dron que sobrevolase la Reserva.

BIBLIOGRAFÍA

Bibliografía

- [1] Assefa M. Melesse, Qihao Weng, Prasad S. Thenkabail and Gabriel B. Senay. Remote Sensing Sensors and Applications in Environmental Resources Mapping and Modelling Sensors.
- [2] Chuvieco, Emilio. Teledetección Ambiental. La observación de la Tierra desde el espacio. 1ª Ed. actualizada, Planeta, S. A., 2010.
- [3] International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Volume XXXVIII, Part 8, Kyoto Japan 2010.
- [4] Gibson, P. y C. H. Power. Introductory Remote Sensing: Digital Image Processing and Applications, London, Routledge, 2000b.
- [5] Eduardo Gómez. Teledetección y sus Aplicaciones. Septiembre 2011.
<http://es.slideshare.net/EDU3364/teledeteccin-9476986?related=1>
Consulta: [15 de Abril de 2016].
- [6] Jensen, J. R. Introductory Digital Image Processing. A Remote Sensing Perspective, Upper Saddle River N.J., Prentice-Hall, 2004.
- [7] Food and Agriculture Organization of the United Nations
<http://www.fao.org/docrep/003/t0446s/T0446S04.htm>
Consulta: [10 de Enero de 2016].
- [8] Lillesand, T. M. y R. W. Kiefer. Remote Sensing and Image Interpretation, New York, John Wiley and Sons, 2000. Referencias 160.
- [9] Mather, P. M. Computer Processing of Remotely Sensed Images, Chichester, John Wiley & Sons, 1998.
- [10] Diario de un radioaficionado
<https://radioaficionado.wordpress.com/2008/07/12/espectro-electromagnetico-desde-el-elf-hasta-rayos-cosmicos/>
Consulta: [9 de Enero de 2016].
- [11] Iván García Cubero. Año Internacional de la Astronomía: El arte en la astronomía

<http://www.wisphysics.es/2009/11/ano-internacional-de-la-astronomia-el-arte-en-la-astronomia>

Consulta: [10 de Enero de 2016].

[12] Short, N. M. The Remote Sensing Tutorial (An Online Handbook). Applied Information Sciences Branch. NASA's Goddard Space Flight Center, 2007.

[13] Centre for Remote Imaging, Sensing & Processing

<http://www.crisp.nus.edu.sg/~research/tutorial/image.htm>

Consulta: [20 de Enero de 2016].

[14] Digital Globe

<https://www.digitalglobe.com/>

Consulta: [15 Octubre de 2014].

[15] Emilio Chuvieco. Fundamentos de Teledetección espacial. Rialp. 3º Edición Revisada, 1996.

[16] Geoimage

<http://www.geoimage.com.au/services/imageprocessing>

Consulta: [15 Octubre de 2014].

[17] Vivone, G., Alparone, L., Chanussot, J., Dalla Mura, M., Garzelli, A., Licciardi, G.A., Restaino, R. and Wald, L., 2015. A critical comparison among pansharpening algorithms. Geoscience and Remote Sensing, IEEE Transactions on, 53(5), pp. 2565-2586.

[18] Amro, I., Mateos, J., Vega, M., Molina, R. and Katsaggelos, A., 2011. A survey of classical methods and new trends in pansharpening of multispectral images. EURASIP Journal on Advances in Signal Processing, 2011(1), pp. 1-22.

[19] Marcello, J., Medina, A. and Eugenio, F., 2013. Evaluation of Spatial and Spectral Effectiveness of Pixel-Level Fusion Techniques. Geoscience and Remote Sensing Letters, IEEE, 10(3), pp. 432-436.

[20] Ehlers, M., Klonus, S., Johan Åstrand, P. and Rosso, P., 2010. Multi-sensor image fusion for pansharpening in remote sensing. International Journal of Image and Data Fusion, 1(1), pp. 25-45.

- [21] Stathaki, T., 2008. Image fusion : algorithms and applications. London [etc.] : (xxk): Academic Press,.
- [22] Li, X., He, M. and Zhang, L., 2013. Hyperspherical color transform based pansharpening method for WorldView-2 satellite images, Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on 2013, IEEE, pp. 520-523.
- [23] Lillo-Saavedra, M., Gonzalo, C., Arquero, A. and Martinez, E., 2005. Fusion of multispectral and panchromatic satellite sensor imagery based on tailored filtering in the Fourier domain. *International Journal of Remote Sensing*, 26(6), pp. 1263-1268.
- [24] Amolins, K., Zhang, Y. and Dare, P., 2007. Wavelet based image fusion techniques — An introduction, review and comparison. *ISPRS Journal of Photogrammetry and Remote Sensing*, 62(4), pp. 249-263.
- [25] Nencini, F., Garzelli, A., Baronti, S. and Alparone, L., 2007. Remote sensing image fusion using the curvelet transform. *Information Fusion*, 8(2), pp. 143-156.
- [26] Zhang, Q. and Guo, B., 2009. Multifocus image fusion using the nonsubsampling contourlet transform. *Signal Processing*, 89(7), pp. 1334-1346.
- [27] Laben, C.A. y Brower, B.V. (1988): Process for Enhancing the Spatial Resolution of Multispectral Imagery Using Pan-Sharpener. Informe técnico, United States Patent 6.011.875.
- [28] Roberti A. Schowengerdt. Remote Sensing. Models and Methods for Image Processing. Elsevier and AP. Third revision.
- [29] Chavez, P.S. An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment* 1988, 24, 459–479.
- [30] Chavez, P.S. Image-Based Atmospheric Corrections. Revisited and Improved. *Photogrammetric Engineering & Remote Sensing* 1996, 62(9), 1025-1036.
- [31] Bernstein, L.S.; Adler-Golden, S.M.; Jin, X.; Gregor B.; Sundberg, R.L. Quick atmospheric correction (QUAC) code for VNIR-SWIR spectral imagery: Algorithm details. In *Proceedings of the IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, Shanghai, China, 4-7 June 2012, pp. 1-4.

- [32] Rahman, H.; Dedieu, G. SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum. *International Journal of Remote Sensing* 1994, 15(1), 123 – 143.
- [33] Staben, G. W.; Pfitzner, K.; Bartolo, R.; Lucieer, A. Empirical line calibration of WorldView-2 satellite imagery to reflectance data: Using quadratic prediction equations. *Remote Sensing Letters* 2012, 3 (6), 521–530.
- [34] Berk, A.; Bernstein, L.S.; Anderson, G.P.; Acharya, P.K.; Robertson, D.C.; Chetwynd, J.H.; Adler-Golden, S.M. MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. *Remote Sensing of Environment* 1998, 65(3), 367-375.
- [35] Adler-Golden, S.M.; Matthew, M.W.; Bernstein, L.S.; Levine, R.Y.; Berk, A.; Richtsmeier, S.C.; Acharya, P.K.; Anderson, G.P.; Felde, G.; Gardner, J.; Hoke, M.; Jeong, L.S.; Pukall, B.; Ratkowski, A.; Burke, H.H. Atmospheric Correction for Short-Wave Spectral Imagery based on MODTRAN4. In *Proceedings SPIE Imaging Spectrometry*, Denver, USA, 18 July 1999, vol. 61.
- [36] Cooley, T.; Anderson, G.P.; Felde, G.W.; Hoke, M.L.; Ratkowski, A.J.; Chetwynd, J.H.; Gardner, J.A.; Adler-Golden, S.M.; Matthew, M.W.; Berk, A.; Bernstein, L.S.; Acharya, P.K.; Miller, D.; Lewis, P. FLAASH, a MODTRAN4-based Atmospheric Correction Algorithm. Its Application and Validation. In *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Toronto, Canada, 24-28 June 2002, vol. 3, pp. 1414–1418.
- [37] Richter, R. A spatially adaptive fast atmospheric correction algorithm. *International Journal of Remote Sensing* 1996, 17(6), 1201-1214.
- [38] Richter, R.; Schläpfer, D. In *Atmospheric/topographic correction for satellite imagery: ATCOR-2/3 User Guide*, DLR report, 2015.
- [39] Vermote, E.; Tanré, D.; Deuzé, J.L.; Herman, M.; Morcrette, J.J.; Kotchenova, S.Y. *Second Simulation of a Satellite Signal in the Solar Spectrum – Vector (6SV)*, 6S User Guide Version 3, 2006. NASA Goddard Space Flight Center, Greenbelt, MD.
- [40] Kotchenova, S.Y.; Vermote, E.F.; Matarrese, R.; Klemm, F.J. Validation of vector version of 6S radiative transfer code for atmospheric correction of satellite data. Part I. Parth radiance. *Applied optics* 2006, 45(26), 6762-6774.

[41] E.F. Vermote and S.Y. Kotchenova, D. Tanre ;J.L. Deuze and M. Herman, Raffaella Matarrese, J.C. Roger, J.J. Morcrette. Modis Land Surface reflectance science computing facility.

<http://6s.ltdri.org/>

Consulta: [10 de Abril de 2016].

[42] Svetlana Y. Kotchenova, Eric F. Vermote, Raffaella Matarrese, and Frank J. Klemm, Jr. Validation of vector version of 6s radiative transfer code for atmospheric correction of satellite data. Part I. Parth radiance.

[43] Eric F. Vermote, Member, IEEE, Didier Tanré, Jean Luc Deuzé, Maurice Herman, and Jean-Jacques Morcrette. Second Simulation of the Satellite Signal in the Solar Spectrum, 6s: An Overview.

[44] LU, D. and WENG, Q. A survey of image classification methods and techniques for improving classification performance, International Journal of Remote Sensing, Vol. 28, No. 5, 10 March, 823–870, 2007.

[45] Schowengerdt, RA. Remote Sensing Models and Methods for Image Processing. Academic Press. 3rd Ed. 2007.

[46] Canty, M. Image Analysis, Classification and Change Detection in Remote Sensing. Taylor and Francis, USA, 2007.

[47] Olaya, Víctor. Sistemas de Información Geográfica. Editor: CreateSpace Independent Publishing Platform. Marzo 2016.

[48] Centro Canadiense de Percepción Remota.

http://www.ccrs.nrcan.gc.ca/resource/tutor/fundam/chapter4/07_e.php

Consulta: [14 de Febrero de 2015].

[49] D. Ballard and C. Brown Computer Vision, Prentice-Hall, Inc., Chap. 6, 1982.

[50] E. Davies Machine Vision: Theory, Algorithms and Practicalities, Academic Press, Chap. 18, 1990.

[51] Richards, J.A. 2013. Remote Sensing digital Image Analysis. Ed. Springer-Verlag.

- [52] Tuia, D., Volpi, M., Copa, L., Kanevski, M., Muñoz-Marí, J. 2011. A survey of active learning algorithms for supervised remote sensing image classification. *IEEE Journal on Selected Topics in Signal Processing*, 5 (3), pp. 606-617
- [53] Tso, B. y Mather, P.M., 2009. *Classification Methods for Remotely Sensed Data*, Second Edition, CRC Press, US.
- [54] Canty, M., 2007. *Image analysis, classification and change detection in remote sensing*. Ed. Taylor & Francis.
- [55] Blaschke, T. 2011. Object based image analysis for remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65, 2-16.
- [56] Lillesand, T.M. and R.W. Kiefer, *Remote Sensing and Image Interpretation*, Ch. 7, *Digital Image Processing*, J. Wiley and Sons, Inc., 2000.
- [57] Marcello Ruiz, Francisco Javier. *Programa de Formación: Introducción al Procesado y Análisis de Imágenes de Teledetección*. Mayo 2013.
- [58] Moik, J.G., *Digital Processing of Remotely Sensed Images*, NASA Special Paper 432, U.S. Govt. Printing Office, 1980.
- [59] Russ, J.C., *The Image Processing Handbook*, CRC Press, 1995.
- [60] Sabins, F.F., *Remote Sensing: Principles and Interpretation*, Ch. 7, *Digital Image Processing*, W.H. Freeman & Co, 1997.
- [61] Vapnik, V. *Estimation of dependences based on empirical data* [in Russian]. Nauka, Moscow. (English translation: Springer-Verlag, New York), 1979.
- [62] Vapnik, V. *The nature of statistical learning theory*. New York: Springer-Verlag, 1995.
- [63] Vapnik, V. *Statistical learning theory*. New York: John Wiley, 1998.
- [64] Burges, C. J. C. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2:121-167, 1998.
- [65] Huang, C., L. S. Davis, and J. R. G. Townshend. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23:725-749, 2002.

- [66] Foody, G. M., and A. Mathur. Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. *Remote Sensing of Environment* 93:107-117, 2004.
- [67] Pal, M., and P. M. Mather. Support vector machines for classification in remote sensing. *International Journal of Remote Sensing* 26:1007-1011, 2005.
- [68] Pal, M. Support vector machine-based feature selection for land cover classification: A case study with DAIS hyperspectral data. *International Journal of Remote Sensing* 27:2877-2894, 2006.
- [69] Cortes, C., and V. Vapnik. Support vector networks. *Machine Learning* 20:273-297, 1995.
- [70] Belousov, A. I., S. A. Verzhakov, and J. Von Frese. A flexible classification approach with optimal generalisation performance: Support vector machines. *Chemometrics and Intelligent Laboratory Systems* 64:15-25, 2002.
- [71] Aizerman, M. A., E. M. Braverman, and L.I. Rozoner. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control* 25:821-837, 1964.
- [72] Boser, B. E., I. M. Guyon, and V. N. Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the 5th Annual Workshop on Computational Learning Theory*, 144-152. Pittsburgh, PA: ACM Press, 1992.
- [73] Halmos, P. R. *A Hilbert space problema book*. Princeton, NJ: D. Van Nostrand Company, Inc., 1967.
- [74] Kolmogorov, A. N., and S.V. Fomin. *Introductory real analysis*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 1970.
- [75] La pastilla roja
<http://lapastillaroja.net/2015/02/ml-algols/>
Consulta: [25 de Febrero de 2015].
- [76] John A. Richards, *Remote Sensing Digital Image Analysis*, fifth edition, Springer-Verlag Berlin Heidelberg, 2013.

- [77] Luntz, A., and V. Brailovsky. On estimation of characters obtained in statistical procedure of recognition. *Technicheskaya Kibernetica*, 3 (in Russian). 1969.
- [78] Jaakkola, T. S., and D. Haussler. Probabilistic kernel regression models. In *Proceedings of the 1999 Conference M on AI and Statistics*. 1999.
- [79] Opper, M., and O. Winter. Gaussian processes and SVM: Mean field and leave-one-out. In *Advances in Large Margin Classifiers*, ed. A. J. Smola, P. L. Bartlett, B. Schölkopf, and D. Schuurmans, 311-326. Cambridge, MA: MIT Press, 2000.
- [80] Vapnik, V., and O. Chapelle. Bounds on error expectation for support vector machines. *Neuronal computation* 12:2013-2036, 2000.
- [81] Chang, C. C., and C. J. Lin. 2007. LIBSVM: A library for support vector machines, 2001. Software available from <http://www.csie.ntu.tw/~cjlin/libsvm>
Consulta: [16 de Marzo de 2016].
- [82] Chapelle, O., V. Vapnik, O. Bousquet, and S. Mukherjee. Choosing multiple parameters for support vector machines. *Machine Learning* 46:131-159, 2002.
- [83] Knerr, S., L. Personnaz, and G. Dreyfus. Single-layer learning revisited: A stepwise procedure for building and training neural network. In *Neurocomputing: Algorithms, architectures and applications*, ed. J. Fogelman, NATO ASI, 41-50. Berlin: Springer-Verlag, 1990.
- [84] Platt, J. C., N. Cristianini, and Shawe-Taylor. Large margin DAGs for multiclass classification. In *Advances in Neural Information Processing Systems*, 547-553. Cambridge, MA: MIT Press, 2000.
- [85] Hsu, C. W., and C. J. Lin. A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks* 13:415-425, 2002.
- [86] Crammer, K., and Y. Singer. On the learnability and design of output codes for multi-class problems. *Machine Learning* 47:201-233, 2002.
- [87] Allan, J.A. Remote sensing of agricultural and forest resources from space. vol. 15, pp. 387-400, 1986.
- [88] Harris Geospatial Solutions
<http://www.harrisgeospatial.com/docs/linearspectralunmixing.html#results>

Consulta: [16 de Febrero de 2016].

[89] Milosz Ciznicki, Krzysztof Kurowski, Antonio Plaza. Graphics processing unit implementation of JPEG2000 for hyperspectral image compression. 2012.

[90] ARTeMISat: Análisis de recursos terrestre y marinos mediante el procesado de imágenes de satélites de alta resolución.

<http://artemisat.ulpgc.es/>

Consulta: [8 de Febrero de 2016].

[91] Gobierno de Canarias. Visor de IDECanarias.

<http://visor.grafcan.es/visorweb/>

Consulta: [23 de Enero de 2016].

[92] Sistema de Información de Ocupación del Suelo en España (SIOSE)

<http://www.siose.es/>

Consulta: [24 de Enero de 2016].

[93] Dr. Antonio Ignacio Hernández Cordero. Tesis Doctoral: Análisis de la vegetación como indicadora de las alteraciones ambientales inducidas por la actividad turística en la Reserva Natural Especial de las Dunas de Maspalomas. 2012.

[94] Dr. Luis F. Hernández Calvento. Tesis Doctoral: Análisis de la evolución del sistema de Dunas de Maspalomas, Gran Canaria, Islas Canarias (1960-2000). 2002.

[95] Sistemas de Información Geográfica, S.A. de C.V.

<http://www.sigsa.info/envi>

Consulta: [16 de Febrero de 2016].

[96] Esri España

<http://www.esri.es/es/>

Consulta: [20 de Septiembre de 2015].

[97] Kavzoglu and Colkesen 2009b; Waske et al. 2010; Huang et al. 2008; Oommen et al. 2008.

[98] Jae H. Min, Young-Chan Lee. An article: Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters.

[99] T. Kavzoglu, I. Colkesen. A kernel functions analysis for support vector machines for land cover classification.

[100] Chih-Wei Hsu, Chih-Chung Chang y Chih-Jen Lin. A practical guide to Support Vector Machine.

[101] Ankur Dixit. Comparison of various models and optimum range of tis parameters used in SVM Classification of Digital Satellite Image.

[102] Xlaojun Yang. Parameterizing Support Vector Machines for Land Cover Classification.

[103] Fauvel, M., J. Chanussot and J.A. Benediktsson, 2006. Evaluation of kernels for multiclass classification of hyperspectral remote sensing data. In: Proceedings, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 06, May 2006, Toulouse France.

<http://fauvel.mathieu.free.fr/publications.html>

Consulta: [14 de Abril de 2016].

[104]Kuemmerle, T., Chaskovskyy, O., Knorn, J., Radeloff, V.C., Kruhlov, I. Forest cover change and illiegal logging in the Ukranian Carpathians in the transition period from 1988 to 2007. Remote Sensing of Environment 113, 1194-1207, 2009.

PLIEGO DE CONDICIONES

C.1. Pliego de condiciones

En el presente capítulo se presenta el pliego de condiciones, donde se detallan los requisitos software y hardware necesarios para la elaboración de este Trabajo Fin de Máster.

A continuación se va a proceder a indicar los requisitos hardware de los equipos, las herramientas software y la imagen satelital utilizada para poder llevar a cabo este Trabajo Fin de Máster.

C.1.1. Requisitos hardware

Para el presente proyecto, se ha utilizado una serie de herramientas hardware que se indican a continuación:

- Ordenador personal Lenovo Z50-70: Ordenador portátil en el que se ha instalado el software ENVI (ENvironment for Visualizing Images) y se ha realizado todo el procesado de las imágenes WorldView 2. Además, se ha redactado la memoria de este Proyecto Final de Carrera.

Las principales características de este PC son:

- Procesador Intel Core i7-4510U a 2,60 GHz.
- 16 GB de memoria RAM.
- 1 TB de disco duro.
- Disco duro extraíble Western Digital de 1 TB: Utilizado para guardar las imágenes procesadas.
- Impresora HP LaserJet 2430 DTN: Utilizada para la impresión del material necesario para el desarrollo del proyecto y de la presente memoria.

C.1.2. Requisitos software

Las herramientas software que fueron necesarias son:

- Sistema Operativo Windows 8.1: Sistema operativo bajo el cual se ha realizado todo el proyecto.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

- Microsoft Office 2010: Conjunto de aplicaciones utilizadas para la elaboración de las tablas, gráficas y diagramas de bloques de la memoria de este Trabajo Fin de Carrera.
- Programa ENVI 5.0: Aplicación informática utilizada para procesar y analizar imágenes geoespaciales instalada en el ordenador portátil para realizar el procesamiento de todas las imágenes WorldView 2.

C.1.3. Imágenes utilizadas

Las imágenes utilizadas para el desarrollo de este Trabajo Fin de Carrera son las imágenes en formato TIFF, del satélite WorldView 2.

- WorldView 2.

Las Palmas de Gran Canaria, 31 de Agosto de 2016

Fdo. Judit Celia Viera Santana

PRESUPUESTO

8.1. Presupuesto

La elaboración del presupuesto del presente Trabajo Fin de Máster tiene en cuenta los siguientes conceptos:

- Honorarios tarifados por tiempo empleado.
- Amortización del software.
- Amortización del hardware.
- Imagen satelital empleada.

Para el cálculo de este presupuesto se ha utilizado la publicación del BOULPGC núm. 11 de Noviembre de 2010 en la cual se especifica la tabla de retribuciones según la categoría para el personal contratado de investigación.

En ella podemos comprobar que la retribución mensual mínima para un investigador en proyecto con la Titulación de Ingeniero y trabajando a tiempo completo (37.5h semanales) asciende a 1.680,59 euros.

8.1.1 Honorarios tarifados por tiempo empleado

Se estima que el desarrollo del Trabajo Fin de Máster se ha requerido un período de 6 meses, a una media de 7,5 horas diarias y cinco días por semana, considerando meses naturales de 30 días, las horas empleadas serían:

$$7,5 \text{ horas/día} * 6 \text{ meses} * 21 \text{ días/mes} = 945 \text{ horas}$$

Anteriormente hemos citado la retribución mensual mínima que ascendía a 1.680,59 euros. Esto se traduce en una tarificación por hora de 10,67 euros. En la Tabla 9.1 podemos apreciar el cálculo de los honorarios para este Trabajo Fin de Máster.

Horas de trabajo	€/hora	Total (€)
945	10,67	10.083,15

Tabla 8.1. Cálculo de honorarios.

8.1.2. Amortización del software

El software empleado para la realización de este proyecto ha sido: ENVI 5.0 y Microsoft Office Word 2010. Se considera un periodo de amortización de 3 años para todos ellos. En la Tabla 9.2 se detallan los cálculos de amortización de software.

Concepto	Coste	Coste/mes	Meses	Total (€)
ENVI 5.0	4.704,30	130,675	6	784,05
MS Office 2010	119	3,31	6	19,86
Total amortización del software				803,91

Tabla 8.2. Amortización del software.

Debemos tener en cuenta que la imagen empleada está disponible de forma gratuita a través de internet y no se incurre en coste alguno en su adquisición.

8.1.3. Amortización del hardware

Para el cálculo de la amortización hardware se considera un periodo de 4 años. El equipo empleado es un ordenador personal con procesador Intel Core i7 (2,60 GHz), con 16 GB de RAM y 1 TB de disco duro. El coste total ascendió a 950 €. La tabla 9.3 detalla los cálculos realizados.

Concepto	Coste	Coste/mes	Meses	Total (€)
Ordenador	950	19,79	6	118,74

Tabla 8.3. Amortización del hardware.

8.1.4. Presupuesto total

El cálculo del presupuesto total se efectúa a partir de la suma de los presupuestos parciales calculados anteriormente. A esta suma habrá que añadirle el costo de la imagen WorldView 2 y el IGIC (Impuesto General Indirecto Canario).

Concepto	Total (€)
Personal	10.083,15
Software	803,91
Hardware	118,74
Imagen WV2	2.180
Coste total	13.185,8

Tabla 8.4. Presupuesto total del Trabajo Fin de Máster.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

El presupuesto total del Trabajo Fin de Máster asciende a TRECE MIL CIENTO OCHENTA Y CINCO CON OCHENTA CÉNTIMOS (13.185,8€).

Las Palmas de Gran Canaria, 31 de Agosto de 2016

Fdo. Judit Celia Viera Santana

ANEXOS

Anexo 1

Separabilidad de las ROIs

Detalladamente hemos incluido en este anexo los pares de separabilidad de cada una de las ROIs empleadas en el proyecto. Diferenciaremos los pares de separabilidad obtenidos para la imagen original y para la imagen fusionada.

- **Pares de separabilidad sobre la imagen original**

- **ROI de entrenamiento**

Input File: 6S_15JUN04114947-M2AS-054281143010_01_P001.TIF
ROI Name: (Jeffries-Matusita, Transformed Divergence)

Arena Seca:

Arena Húmeda: (1.79692999 1.89971571)
Costra Salina: (1.97869653 1.99990111)
Charca: (2.00000000 2.00000000)
Callaos: (1.76698397 1.99657190)
Aluviales: (1.69186444 1.99950441)
Tamarix canariensis: (1.99936134 2.00000000)
Traganum moquini: (1.99996431 2.00000000)
Juncus acutus: (1.99999800 2.00000000)
Suaeda mollis: (1.99957153 1.99999557)
Tetraena fontanesii: (1.87509021 1.99975444)
Launaea arborescens: (1.99958000 2.00000000)

Costra Salina:

Arena Seca: (1.97869653 1.99990111)
Arena Húmeda: (1.99768885 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99738847 2.00000000)
Aluviales: (1.98879544 2.00000000)
Tamarix canariensis: (1.99992705 2.00000000)
Traganum moquini: (1.99994644 2.00000000)
Juncus acutus: (1.99999957 2.00000000)
Suaeda mollis: (1.99998302 2.00000000)
Tetraena fontanesii: (1.99501767 1.99999999)
Launaea arborescens: (1.99999805 2.00000000)

Arena Húmeda:

Arena Seca: (1.79692999 1.89971571)
Costra Salina: (1.99768885 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.88963152 1.97804905)
Aluviales: (1.91798155 1.98451591)
Tamarix canariensis: (1.99885896 2.00000000)
Traganum moquini: (1.99999991 2.00000000)
Juncus acutus: (1.99998262 2.00000000)
Suaeda mollis: (1.99999484 1.99999943)
Tetraena fontanesii: (1.73121510 1.99977563)
Launaea arborescens: (1.99999986 2.00000000)

Charca:

Arena Seca: (2.00000000 2.00000000)
Arena Húmeda: (2.00000000 2.00000000)
Costra Salina: (2.00000000 2.00000000)
Callaos: (2.00000000 2.00000000)
Aluviales: (2.00000000 2.00000000)
Tamarix canariensis: (2.00000000 2.00000000)
Traganum moquini: (2.00000000 2.00000000)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (2.00000000 2.00000000)
Tetraena fontanesii: (2.00000000 2.00000000)
Launaea arborescens: (2.00000000 2.00000000)

Callaos:

Arena Seca: (1.76698397 1.99657190)
Arena Húmeda: (1.88963152 1.97804905)
Costra Salina: (1.99738847 2.00000000)
Charca: (2.00000000 2.00000000)
Aluviales: (1.93496891 1.99916969)
Tamarix canariensis: (1.99969954 2.00000000)
Traganum moquini: (1.99999745 2.00000000)
Juncus acutus: (1.99999981 2.00000000)
Suaeda mollis: (1.99999981 2.00000000)
Tetraena fontanesii: (1.92746771 1.99999999)
Launaea arborescens: (2.00000000 2.00000000)

Tamarix canariensis:

Arena Seca: (1.99936134 2.00000000)
Arena Húmeda: (1.99885896 2.00000000)
Costra Salina: (1.99992705 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99969954 2.00000000)
Aluviales: (1.99986330 2.00000000)
Traganum moquini: (1.78973581 1.98056511)
Juncus acutus: (1.96576522 1.99970681)
Suaeda mollis: (1.91154564 2.00000000)
Tetraena fontanesii: (1.89393016 1.99992625)
Launaea arborescens: (1.98976335 2.00000000)

Aluviales:

Arena Seca: (1.69186444 1.99950441)
Arena Húmeda: (1.91798155 1.98451591)
Costra Salina: (1.98879544 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.93496891 1.99916969)
Tamarix canariensis: (1.99986330 2.00000000)
Traganum moquini: (1.99999910 2.00000000)
Juncus acutus: (1.99999999 2.00000000)
Suaeda mollis: (1.99999515 2.00000000)
Tetraena fontanesii: (1.95127174 2.00000000)
Launaea arborescens: (1.99999922 2.00000000)

Traganum moquini:

Arena Seca: (1.99996431 2.00000000)
Arena Húmeda: (1.99999991 2.00000000)
Costra Salina: (1.99994644 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99999745 2.00000000)
Aluviales: (1.99999910 2.00000000)
Tamarix canariensis: (1.78973581 1.98056511)
Juncus acutus: (1.99915136 1.99999988)
Suaeda mollis: (1.91615809 1.99848903)
Tetraena fontanesii: (1.81840871 1.98338034)
Launaea arborescens: (1.92590209 1.99996115)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

```
Juncus acutus:
  Arena Seca: (1.99999800 2.00000000)
  Arena Húmeda: (1.99998262 2.00000000)
  Costra Salina: (1.99999957 2.00000000)
  Charca: (2.00000000 2.00000000)
  Callaos: (1.99999981 2.00000000)
  Aluviales: (1.99999999 2.00000000)
  Tamarix canariensis: (1.96576522 1.99970681)
  Traganum moquini: (1.99915136 1.99999988)
  Suaeda mollis: (1.99833803 2.00000000)
  Tetraena fontanesii: (1.98722851 1.99929264)
  Launaea arborescens: (1.99986567 2.00000000)

Tetraena fontanesii:
  Arena Seca: (1.87509021 1.99975444)
  Arena Húmeda: (1.73121510 1.99977563)
  Costra Salina: (1.99501767 1.99999999)
  Charca: (2.00000000 2.00000000)
  Callaos: (1.92746771 1.99999999)
  Aluviales: (1.95127174 2.00000000)
  Tamarix canariensis: (1.89393016 1.99992625)
  Traganum moquini: (1.81840871 1.98338034)
  Juncus acutus: (1.98722851 1.99929264)
  Suaeda mollis: (1.35540849 1.79964447)
  Launaea arborescens: (1.97492336 1.99999999)

Suaeda mollis:
  Arena Seca: (1.99957153 1.99999557)
  Arena Húmeda: (1.99999484 1.99999943)
  Costra Salina: (1.99998302 2.00000000)
  Charca: (2.00000000 2.00000000)
  Callaos: (1.99999981 2.00000000)
  Aluviales: (1.99999515 2.00000000)
  Tamarix canariensis: (1.91154564 2.00000000)
  Traganum moquini: (1.91615809 1.99848903)
  Juncus acutus: (1.99833803 2.00000000)
  Tetraena fontanesii: (1.35540849 1.79964447)
  Launaea arborescens: (1.98598959 2.00000000)

Launaea arborescens:
  Arena Seca: (1.99958000 2.00000000)
  Arena Húmeda: (1.99999986 2.00000000)
  Costra Salina: (1.99999805 2.00000000)
  Charca: (2.00000000 2.00000000)
  Callaos: (2.00000000 2.00000000)
  Aluviales: (1.99999922 2.00000000)
  Tamarix canariensis: (1.98976335 2.00000000)
  Traganum moquini: (1.92590209 1.99996115)
  Juncus acutus: (1.99986567 2.00000000)
  Suaeda mollis: (1.98598959 2.00000000)
  Tetraena fontanesii: (1.97492336 1.99999999)
```

Figura A1.1. Pares de separabilidad para la ROI de entrenamiento sobre la imagen original.

○ ROI de evaluación

Hacemos lo mismo que para la ROI de entrenamiento, plasmamos en la siguiente figura los pares de separabilidad para la ROI de evaluación sobre la imagen original.

```
Input File: 6S_15JUN04114947-M2AS-054281143010_01_P001.TIF
ROI Name: (Jeffries-Matusita, Transformed Divergence)

Arena Seca:
  Arena Húmeda: (1.85479848 1.99545495)
  Charca: (2.00000000 2.00000000)
  Callaos: (1.74447327 1.99992597)
  Aluviales: (1.76562091 1.99990972)
  Tamarix canariensis: (1.99999996 2.00000000)
  Traganum moquini: (1.99994740 2.00000000)
  Juncus acutus: (2.00000000 2.00000000)
  Suaeda mollis: (1.97524042 1.99988781)
  Tetraena fontanesii: (1.99613994 2.00000000)
  Launaea arborescens: (1.99998624 2.00000000)
  Costra Salina: (1.95567833 1.99999798)

Charca:
  Arena Seca: (2.00000000 2.00000000)
  Arena Húmeda: (2.00000000 2.00000000)
  Callaos: (2.00000000 2.00000000)
  Aluviales: (2.00000000 2.00000000)
  Tamarix canariensis: (2.00000000 2.00000000)
  Traganum moquini: (2.00000000 2.00000000)
  Juncus acutus: (2.00000000 2.00000000)
  Suaeda mollis: (2.00000000 2.00000000)
  Tetraena fontanesii: (2.00000000 2.00000000)
  Launaea arborescens: (2.00000000 2.00000000)
  Costra Salina: (2.00000000 2.00000000)

Arena Húmeda:
  Arena Seca: (1.85479848 1.99545495)
  Charca: (2.00000000 2.00000000)
  Callaos: (1.84171191 1.93217580)
  Aluviales: (1.87336615 1.99428411)
  Tamarix canariensis: (1.99999781 2.00000000)
  Traganum moquini: (1.99999025 2.00000000)
  Juncus acutus: (2.00000000 2.00000000)
  Suaeda mollis: (1.99230928 1.99982151)
  Tetraena fontanesii: (1.97624501 2.00000000)
  Launaea arborescens: (1.99999999 2.00000000)
  Costra Salina: (1.99504026 1.99999998)

Callaos:
  Arena Seca: (1.74447327 1.99992597)
  Arena Húmeda: (1.84171191 1.93217580)
  Charca: (2.00000000 2.00000000)
  Aluviales: (1.93274383 1.99990189)
  Tamarix canariensis: (2.00000000 2.00000000)
  Traganum moquini: (1.99996316 2.00000000)
  Juncus acutus: (2.00000000 2.00000000)
  Suaeda mollis: (1.99796743 1.99999282)
  Tetraena fontanesii: (1.98336460 2.00000000)
  Launaea arborescens: (2.00000000 2.00000000)
  Costra Salina: (1.99616696 2.00000000)
```

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Aluviales: Arena Seca: (1.76562091 1.99990972) Arena Húmeda: (1.87336615 1.99428411) Charca: (2.00000000 2.00000000) Callaos: (1.93274383 1.99990189) Tamarix canariensis: (2.00000000 2.00000000) Traganum moquini: (1.99970919 2.00000000) Juncus acutus: (2.00000000 2.00000000) Suaeda mollis: (1.97513204 1.99999991) Tetraena fontanesii: (1.97449867 2.00000000) Launaea arborescens: (1.99999988 2.00000000) Costra Salina: (1.97884557 2.00000000)	Traganum moquini: Arena Seca: (1.99994740 2.00000000) Arena Húmeda: (1.99999025 2.00000000) Charca: (2.00000000 2.00000000) Callaos: (1.99996316 2.00000000) Aluviales: (1.99970919 2.00000000) Tamarix canariensis: (1.87577540 1.91266827) Juncus acutus: (1.99985020 1.99999824) Suaeda mollis: (1.86220280 1.96992328) Tetraena fontanesii: (1.78496136 1.99795051) Launaea arborescens: (1.88909579 1.99884178) Costra Salina: (1.99996783 2.00000000)
Tamarix canariensis: Arena Seca: (1.99999996 2.00000000) Arena Húmeda: (1.99999781 2.00000000) Charca: (2.00000000 2.00000000) Callaos: (2.00000000 2.00000000) Aluviales: (2.00000000 2.00000000) Traganum moquini: (1.87577540 1.91266827) Juncus acutus: (1.97499990 1.99844658) Suaeda mollis: (1.99888088 1.99999629) Tetraena fontanesii: (1.98854816 1.99999973) Launaea arborescens: (1.99828650 1.99999521) Costra Salina: (1.99999999 2.00000000)	Juncus acutus: Arena Seca: (2.00000000 2.00000000) Arena Húmeda: (2.00000000 2.00000000) Charca: (2.00000000 2.00000000) Callaos: (2.00000000 2.00000000) Aluviales: (2.00000000 2.00000000) Tamarix canariensis: (1.97499990 1.99844658) Traganum moquini: (1.99985020 1.99999824) Suaeda mollis: (1.99996045 2.00000000) Tetraena fontanesii: (1.99935358 1.99999998) Launaea arborescens: (1.99943381 2.00000000) Costra Salina: (2.00000000 2.00000000)
Suaeda mollis: Arena Seca: (1.97524042 1.99988781) Arena Húmeda: (1.99230928 1.99982151) Charca: (2.00000000 2.00000000) Callaos: (1.99796743 1.99999282) Aluviales: (1.97513204 1.99999991) Tamarix canariensis: (1.99888088 1.99999629) Traganum moquini: (1.86220280 1.96992328) Juncus acutus: (1.99996045 2.00000000) Tetraena fontanesii: (1.69048457 1.95321039) Launaea arborescens: (1.98619685 1.99999680) Costra Salina: (1.99509338 2.00000000)	Launaea arborescens: Arena Seca: (1.99998624 2.00000000) Arena Húmeda: (1.99999999 2.00000000) Charca: (2.00000000 2.00000000) Callaos: (2.00000000 2.00000000) Aluviales: (1.99999988 2.00000000) Tamarix canariensis: (1.99828650 1.99999521) Traganum moquini: (1.88909579 1.99884178) Juncus acutus: (1.99943381 2.00000000) Suaeda mollis: (1.98619685 1.99999680) Tetraena fontanesii: (1.97067724 2.00000000) Costra Salina: (1.99998811 1.99999999)
Tetraena fontanesii: Arena Seca: (1.99613994 2.00000000) Arena Húmeda: (1.97624501 2.00000000) Charca: (2.00000000 2.00000000) Callaos: (1.98336460 2.00000000) Aluviales: (1.97449867 2.00000000) Tamarix canariensis: (1.98854816 1.99999973) Traganum moquini: (1.78496136 1.99795051) Juncus acutus: (1.99935358 1.99999998) Suaeda mollis: (1.69048457 1.95321039) Launaea arborescens: (1.97067724 2.00000000) Costra Salina: (1.99904347 2.00000000)	Costra Salina: Arena Seca: (1.95567833 1.99999798) Arena Húmeda: (1.99504026 1.99999998) Charca: (2.00000000 2.00000000) Callaos: (1.99616696 2.00000000) Aluviales: (1.97884557 2.00000000) Tamarix canariensis: (1.99999999 2.00000000) Traganum moquini: (1.99996783 2.00000000) Juncus acutus: (2.00000000 2.00000000) Suaeda mollis: (1.99509338 2.00000000) Tetraena fontanesii: (1.99904347 2.00000000) Launaea arborescens: (1.99998811 1.99999999)

Figura A1.2. Pares de separabilidad para la ROI de evaluación sobre la imagen original.

- **Pares de separabilidad sobre la imagen fusionada**

Siguiente el procedimiento desarrollado para la imagen original, repetimos el proceso pero empleando la imagen fusionada. Se exponen los pares de separabilidad obtenidos para las ROIs de entrenamiento y evaluación sobre la imagen fusionada.

- **ROI de entrenamiento**

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Input File: GS_6S_15JUN04114947-M2AS-054281143010_01_P001.tif
ROI Name: (Jeffries-Matusita, Transformed Divergence)

Arena Seca:

Arena Húmeda: (1.86262674 1.94826450)
Costra Salina: (1.97280700 1.99999246)
Charca: (2.00000000 2.00000000)
Callaos: (1.80736697 1.99561123)
Aluviales: (1.79282710 1.99997742)
Tamarix canariensis: (1.99985959 2.00000000)
Tragnum moquini: (1.99999583 2.00000000)
Juncus acutus: (1.99999995 2.00000000)
Suaeda mollis: (1.99744678 1.99997986)
Tetraena fontanesii: (1.89522739 1.99994164)
Launaea arborescens: (1.99987520 2.00000000)

Arena Húmeda:

Arena Seca: (1.86262674 1.94826450)
Costra Salina: (1.99169803 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.91280670 1.98259022)
Aluviales: (1.93827004 1.97445191)
Tamarix canariensis: (1.99977523 2.00000000)
Tragnum moquini: (1.99999999 2.00000000)
Juncus acutus: (1.99999956 2.00000000)
Suaeda mollis: (1.99723260 1.99998774)
Tetraena fontanesii: (1.75162635 1.99999818)
Launaea arborescens: (1.99999946 2.00000000)

Callaos:

Arena Seca: (1.80736697 1.99561123)
Arena Húmeda: (1.91280670 1.98259022)
Costra Salina: (1.98995986 2.00000000)
Charca: (2.00000000 2.00000000)
Aluviales: (1.93947026 1.99988535)
Tamarix canariensis: (1.99995461 2.00000000)
Tragnum moquini: (1.99999997 2.00000000)
Juncus acutus: (1.99999999 2.00000000)
Suaeda mollis: (1.99996615 2.00000000)
Tetraena fontanesii: (1.96209837 2.00000000)
Launaea arborescens: (1.99999999 2.00000000)

Aluviales:

Arena Seca: (1.79282710 1.99997742)
Arena Húmeda: (1.93827004 1.97445191)
Costra Salina: (1.98193629 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.93947026 1.99988535)
Tamarix canariensis: (1.99996865 2.00000000)
Tragnum moquini: (2.00000000 2.00000000)
Juncus acutus: (1.99999999 2.00000000)
Suaeda mollis: (1.99993166 2.00000000)
Tetraena fontanesii: (1.94057375 2.00000000)
Launaea arborescens: (1.99999999 2.00000000)

Juncus acutus:

Arena Seca: (1.99999995 2.00000000)
Arena Húmeda: (1.99999956 2.00000000)
Costra Salina: (1.99999962 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99999999 2.00000000)
Aluviales: (1.99999999 2.00000000)
Tamarix canariensis: (1.99186692 1.99999353)
Tragnum moquini: (1.99965553 2.00000000)
Suaeda mollis: (1.99963752 2.00000000)
Tetraena fontanesii: (1.99729752 1.99999559)
Launaea arborescens: (1.99884642 2.00000000)

Suaeda mollis:

Arena Seca: (1.99744678 1.99997986)
Arena Húmeda: (1.99723260 1.99998774)
Costra Salina: (1.99986861 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99996615 2.00000000)
Aluviales: (1.99993166 2.00000000)
Tamarix canariensis: (1.93755221 2.00000000)
Tragnum moquini: (1.93971129 1.99948661)
Juncus acutus: (1.99963752 2.00000000)
Tetraena fontanesii: (1.42083267 1.84945612)
Launaea arborescens: (1.95511284 1.99999999)

Costra Salina:

Arena Seca: (1.97280700 1.99999246)
Arena Húmeda: (1.99169803 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.98995986 2.00000000)
Aluviales: (1.98193629 2.00000000)
Tamarix canariensis: (1.99992777 2.00000000)
Tragnum moquini: (1.99997929 2.00000000)
Juncus acutus: (1.99999962 2.00000000)
Suaeda mollis: (1.99986861 2.00000000)
Tetraena fontanesii: (1.98745289 1.99999938)
Launaea arborescens: (1.99999877 2.00000000)

Charca:

Arena Seca: (2.00000000 2.00000000)
Arena Húmeda: (2.00000000 2.00000000)
Costra Salina: (2.00000000 2.00000000)
Callaos: (2.00000000 2.00000000)
Aluviales: (2.00000000 2.00000000)
Tamarix canariensis: (2.00000000 2.00000000)
Tragnum moquini: (2.00000000 2.00000000)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (2.00000000 2.00000000)
Tetraena fontanesii: (2.00000000 2.00000000)
Launaea arborescens: (2.00000000 2.00000000)

Tamarix canariensis:

Arena Seca: (1.99985959 2.00000000)
Arena Húmeda: (1.99977523 2.00000000)
Costra Salina: (1.99992777 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99995461 2.00000000)
Aluviales: (1.99996865 2.00000000)
Tragnum moquini: (1.83216484 1.99726198)
Juncus acutus: (1.99186692 1.99999353)
Suaeda mollis: (1.93755221 2.00000000)
Tetraena fontanesii: (1.95535903 1.99996801)
Launaea arborescens: (1.98341384 2.00000000)

Tragnum moquini:

Arena Seca: (1.99999583 2.00000000)
Arena Húmeda: (1.99999999 2.00000000)
Costra Salina: (1.99997929 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99999997 2.00000000)
Aluviales: (2.00000000 2.00000000)
Tamarix canariensis: (1.83216484 1.99726198)
Juncus acutus: (1.99965553 2.00000000)
Suaeda mollis: (1.93971129 1.99948661)
Tetraena fontanesii: (1.91092272 1.99711404)
Launaea arborescens: (1.94447090 1.99999800)

Tetraena fontanesii:

Arena Seca: (1.89522739 1.99994164)
Arena Húmeda: (1.75162635 1.99999818)
Costra Salina: (1.98745289 1.99999938)
Charca: (2.00000000 2.00000000)
Callaos: (1.96209837 2.00000000)
Aluviales: (1.94057375 2.00000000)
Tamarix canariensis: (1.95535903 1.99996801)
Tragnum moquini: (1.91092272 1.99711404)
Juncus acutus: (1.99729752 1.99999559)
Suaeda mollis: (1.42083267 1.84945612)
Launaea arborescens: (1.93767019 2.00000000)

Launaea arborescens:

Arena Seca: (1.99987520 2.00000000)
Arena Húmeda: (1.99999946 2.00000000)
Costra Salina: (1.99999877 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99999999 2.00000000)
Aluviales: (1.99999999 2.00000000)
Tamarix canariensis: (1.98341384 2.00000000)
Tragnum moquini: (1.94447090 1.99999800)
Juncus acutus: (1.99884642 2.00000000)
Suaeda mollis: (1.95511284 1.99999999)
Tetraena fontanesii: (1.93767019 2.00000000)

Figura A1.3. Pares de separabilidad para la ROI de entrenamiento sobre la imagen fusionada.

○ ROI de evaluación

Finalmente, se muestran en la Figura A1.4 los pares de separabilidad de la ROI de evaluación sobre la imagen fusionada.

Input File: GS_6S_15JUN04114947-M2AS-054281143010_01_P001.tif
ROI Name: (Jeffries-Matusita, Transformed Divergence)

Arena Seca:

Arena Húmeda: (1.90638974 1.99925979)
Charca: (2.00000000 2.00000000)
Callaos: (1.82980136 1.99975491)
Aluviales: (1.79844089 1.99986457)
Tamarix canariensis: (1.99999996 2.00000000)
Tragnum moquini: (1.99977857 2.00000000)
Juncus acutus: (1.99999995 2.00000000)
Suaeda mollis: (1.98401613 1.99945209)
Tetraena fontanesii: (1.99827489 2.00000000)
Launaea arborescens: (1.99996980 2.00000000)
Costra Salina: (1.96361276 1.99999933)

Arena Húmeda:

Arena Seca: (1.90638974 1.99925979)
Charca: (2.00000000 2.00000000)
Callaos: (1.84557622 1.94392596)
Aluviales: (1.86584592 1.97802028)
Tamarix canariensis: (1.99999925 2.00000000)
Tragnum moquini: (1.99994574 2.00000000)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (1.98441983 1.99928785)
Tetraena fontanesii: (1.98869840 2.00000000)
Launaea arborescens: (1.99999240 2.00000000)
Costra Salina: (1.98825468 2.00000000)

Aluviales:

Arena Seca: (1.79844089 1.99986457)
Arena Húmeda: (1.86584592 1.97802028)
Charca: (2.00000000 2.00000000)
Callaos: (1.94062207 1.99997176)
Tamarix canariensis: (1.99999999 2.00000000)
Tragnum moquini: (1.99891786 2.00000000)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (1.99590534 1.99999997)
Tetraena fontanesii: (1.98260431 2.00000000)
Launaea arborescens: (2.00000000 2.00000000)
Costra Salina: (1.97587689 2.00000000)

Tamarix canariensis:

Arena Seca: (1.99999996 2.00000000)
Arena Húmeda: (1.99999925 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (2.00000000 2.00000000)
Aluviales: (1.99999999 2.00000000)
Tragnum moquini: (1.81746853 1.89802351)
Juncus acutus: (1.98629619 1.99971919)
Suaeda mollis: (1.99805157 2.00000000)
Tetraena fontanesii: (1.97808022 1.99999483)
Launaea arborescens: (1.98839549 1.99999966)
Costra Salina: (1.99999353 2.00000000)

Charca:

Arena Seca: (2.00000000 2.00000000)
Arena Húmeda: (2.00000000 2.00000000)
Callaos: (2.00000000 2.00000000)
Aluviales: (2.00000000 2.00000000)
Tamarix canariensis: (2.00000000 2.00000000)
Tragnum moquini: (2.00000000 2.00000000)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (2.00000000 2.00000000)
Tetraena fontanesii: (2.00000000 2.00000000)
Launaea arborescens: (2.00000000 2.00000000)
Costra Salina: (2.00000000 2.00000000)

Callaos:

Arena Seca: (1.82980136 1.99975491)
Arena Húmeda: (1.84557622 1.94392596)
Charca: (2.00000000 2.00000000)
Aluviales: (1.94062207 1.99997176)
Tamarix canariensis: (2.00000000 2.00000000)
Tragnum moquini: (1.99979051 2.00000000)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (1.99974618 1.99999965)
Tetraena fontanesii: (1.99785185 2.00000000)
Launaea arborescens: (1.99999997 2.00000000)
Costra Salina: (1.98266180 2.00000000)

Tragnum moquini:

Arena Seca: (1.99977857 2.00000000)
Arena Húmeda: (1.99994574 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99979051 2.00000000)
Aluviales: (1.99891786 2.00000000)
Tamarix canariensis: (1.81746853 1.89802351)
Juncus acutus: (1.99854706 1.99990869)
Suaeda mollis: (1.94127648 1.99804323)
Tetraena fontanesii: (1.81885289 1.97767150)
Launaea arborescens: (1.87495934 1.99870151)
Costra Salina: (1.99787688 1.99999715)

Juncus acutus:

Arena Seca: (1.99999995 2.00000000)
Arena Húmeda: (2.00000000 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (2.00000000 2.00000000)
Aluviales: (2.00000000 2.00000000)
Tamarix canariensis: (1.98629619 1.99971919)
Tragnum moquini: (1.99854706 1.99990869)
Suaeda mollis: (1.99988417 2.00000000)
Tetraena fontanesii: (1.99593076 1.99999829)
Launaea arborescens: (1.99679653 2.00000000)
Costra Salina: (2.00000000 2.00000000)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Suaeda mollis:
Arena Seca: (1.98601613 1.99945209)
Arena Húmeda: (1.98441983 1.99928785)
Charca: (2.00000000 2.00000000)
Callaos: (1.99974618 1.99999965)
Aluviales: (1.99590534 1.99999997)
Tamarix canariensis: (1.99805157 2.00000000)
Traganum moquini: (1.94127648 1.99804323)
Juncus acutus: (1.99988417 2.00000000)
Tetraena fontanesii: (1.68712230 1.97012929)
Launaea arborescens: (1.92831188 1.99958088)
Costra Salina: (1.98447534 1.99992689)

Tetraena fontanesii:
Arena Seca: (1.99827489 2.00000000)
Arena Húmeda: (1.98869840 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99785185 2.00000000)
Aluviales: (1.98260431 2.00000000)
Tamarix canariensis: (1.97808022 1.99999483)
Traganum moquini: (1.81885289 1.97767150)
Juncus acutus: (1.99593076 1.99999829)
Suaeda mollis: (1.68712230 1.97012929)
Launaea arborescens: (1.92768733 2.00000000)
Costra Salina: (1.99285459 1.99999971)

Launaea arborescens:
Arena Seca: (1.99996980 2.00000000)
Arena Húmeda: (1.99999240 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.99999997 2.00000000)
Aluviales: (2.00000000 2.00000000)
Tamarix canariensis: (1.98839549 1.99999966)
Traganum moquini: (1.87495934 1.99870151)
Juncus acutus: (1.99679653 2.00000000)
Suaeda mollis: (1.92831188 1.99958088)
Tetraena fontanesii: (1.92768733 2.00000000)
Costra Salina: (1.99911849 1.99999989)

Costra Salina:
Arena Seca: (1.96361276 1.99999933)
Arena Húmeda: (1.98825468 2.00000000)
Charca: (2.00000000 2.00000000)
Callaos: (1.98266180 2.00000000)
Aluviales: (1.97587689 2.00000000)
Tamarix canariensis: (1.99999353 2.00000000)
Traganum moquini: (1.99787688 1.99999715)
Juncus acutus: (2.00000000 2.00000000)
Suaeda mollis: (1.98447534 1.99992689)
Tetraena fontanesii: (1.99285459 1.99999971)
Launaea arborescens: (1.99911849 1.99999989)

Figura A1.4. Pares de separabilidad para la ROI de evaluación sobre la imagen fusionada.

Anexo 2

Matrices de Confusión

Este anexo está formado por las matrices de confusión obtenidas tras evaluar las clasificaciones realizadas a lo largo de este Trabajo Fin de Máster. Dichas matrices nos permiten evaluar los conflictos que se producen entre las clases.

Expondremos los resultados obtenidos tras clasificar la Reserva Natural de las Dunas de Maspalomas. Podremos ver las matrices de confusión obtenidas para cada clasificador sobre la imagen original y posteriormente los resultados sobre la imagen fusionada.

- **Imagen Original**

Maximum Likelihood

Class	Ground Truth (Pixels)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0	0	0	0	0	0
Arena Seca	1358	10	0	0	0	3
Arena Húmeda	4	901	0	0	0	1
Costra Salina	35	1	38	0	0	1
Charca	0	0	0	423	0	0
Callaos	0	1	0	0	0	87
Aluviales	43	44	0	0	0	0
Tamarix canar	0	0	0	0	0	0
Traganum moqu	0	0	0	0	0	0
Juncus acutus	0	0	0	0	0	0
Suaeda mollis	0	0	0	0	0	0
Tetraena font	17	28	0	0	0	2
Launaea arbor	4	0	2	0	0	0
Total	1461	985	40	423	94	

Class	Ground Truth (Pixels)					
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis	
Unclassified	0	0	0	0	0	0
Arena Seca	15	0	0	0	0	3
Arena Húmeda	2	0	0	0	0	0
Costra Salina	0	0	0	0	0	0
Charca	0	0	0	0	0	0
Callaos	0	0	0	0	0	0
Aluviales	97	0	0	0	0	0
Tamarix canar	0	206	22	6	0	0
Traganum moqu	0	0	127	0	0	1
Juncus acutus	0	1	0	37	0	0
Suaeda mollis	0	0	0	0	20	0
Tetraena font	5	0	15	1	0	14
Launaea arbor	0	5	6	1	0	2
Total	119	212	170	45	40	

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	1389
Arena Húmeda	0	0	908
Costra Salina	0	0	75
Charca	0	0	423
Callaos	0	0	88
Aluviales	0	0	184
Tamarix canar	0	1	235
Traganum moqu	2	6	136
Juncus acutus	0	0	38
Suaeda mollis	4	0	24
Tetraena font	8	1	91
Launaea arbor	0	40	60
Total	14	48	3651

(a)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	92.95	1.02	0.00	0.00	3.19
Arena Húmeda	0.27	91.47	0.00	0.00	1.06
Costra Salina	2.40	0.10	95.00	0.00	1.06
Charca	0.00	0.00	0.00	100.00	0.00
Callaos	0.00	0.10	0.00	0.00	92.55
Aluviales	2.94	4.47	0.00	0.00	0.00
Tamarix canar	0.00	0.00	0.00	0.00	0.00
Traganum moqu	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	0.00
Tetraena font	1.16	2.84	0.00	0.00	2.13
Launaea arbor	0.27	0.00	5.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	12.61	0.00	0.00	0.00	7.50
Arena Húmeda	1.68	0.00	0.00	0.00	0.00
Costra Salina	0.00	0.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00
Callaos	0.00	0.00	0.00	0.00	0.00
Aluviales	81.51	0.00	0.00	0.00	0.00
Tamarix canar	0.00	97.17	12.94	13.33	0.00
Traganum moqu	0.00	0.00	74.71	0.00	2.50
Juncus acutus	0.00	0.47	0.00	82.22	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	50.00
Tetraena font	4.20	0.00	8.82	2.22	35.00
Launaea arbor	0.00	2.36	3.53	2.22	5.00
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)		Total
	Tetraena font	Launaea arbor	
Unclassified	0.00	0.00	0.00
Arena Seca	0.00	0.00	38.04
Arena Húmeda	0.00	0.00	24.87
Costra Salina	0.00	0.00	2.05
Charca	0.00	0.00	11.59
Callaos	0.00	0.00	2.41
Aluviales	0.00	0.00	5.04
Tamarix canar	0.00	2.08	6.44
Traganum moqu	14.29	12.50	3.73
Juncus acutus	0.00	0.00	1.04
Suaeda mollis	28.57	0.00	0.66
Tetraena font	57.14	2.08	2.49
Launaea arbor	0.00	83.33	1.64
Total	100.00	100.00	100.00

(b)

Class	Commission		Omission	
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	2.23	7.05	31/1389	103/1461
Arena Húmeda	0.77	8.53	7/908	84/985
Costra Salina	49.33	5.00	37/75	2/40
Charca	0.00	0.00	0/423	0/423
Callaos	1.14	7.45	1/88	7/94
Aluviales	47.28	18.49	87/184	22/119
Tamarix canar	12.34	2.83	29/235	6/212
Traganum moqu	6.62	25.29	9/136	43/170
Juncus acutus	2.63	17.78	1/38	8/45
Suaeda mollis	16.67	50.00	4/24	20/40
Tetraena font	91.21	42.86	83/91	6/14
Launaea arbor	33.33	16.67	20/60	8/48

(c)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Arena Seca	92.95	97.77	1358/1461	1358/1389
Arena Húmeda	91.47	99.23	901/985	901/908
Costra Salina	95.00	50.67	38/40	38/75
Charca	100.00	100.00	423/423	423/423
Callaos	92.55	98.86	87/94	87/88
Aluviales	81.51	52.72	97/119	97/184
Tamarix canar	97.17	87.66	206/212	206/235
Traganum moqu	74.71	93.38	127/170	127/136
Juncus acutus	82.22	97.37	37/45	37/38
Suaeda mollis	50.00	83.33	20/40	20/24
Tetraena font	57.14	8.79	8/14	8/91
Launaea arbor	83.33	66.67	40/48	40/60

(d)

**Tabla A2.1. Matriz de Confusión para el clasificador *Maximum Likelihood* sobre imagen original:
(a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commision y Omission, (d) Prod. Acc.
y User Acc.**

Class	Ground Truth (Pixels)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0	0	0	0	0	0
Arena Seca	1380	3	0	0	0	3
Arena Húmeda	0	933	0	0	0	0
Costra Salina	40	0	39	0	0	0
Charca	0	0	0	423	0	0
Callaos	0	0	0	0	90	0
Aluviales	33	48	0	0	0	0
Tamarix canar	0	0	0	0	0	0
Traganum moqu	0	0	0	0	0	0
Juncus acutus	0	0	0	0	0	0
Suaeda mollis	0	0	0	0	0	0
Tetraena font	8	1	0	0	0	1
Launaea arbor	0	0	1	0	0	0
Total	1461	985	40	423	94	

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	4	1	1	0	0
Arena Húmeda	0	0	0	0	0
Costra Salina	0	0	0	0	0
Charca	0	0	0	0	0
Callaos	0	0	0	0	0
Aluviales	109	0	0	0	0
Tamarix canar	0	207	12	7	0
Traganum moqu	0	0	149	0	0
Juncus acutus	0	0	0	37	0
Suaeda mollis	0	0	0	0	22
Tetraena font	5	0	6	0	18
Launaea arbor	1	4	2	1	0
Total	119	212	170	45	40

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	1392
Arena Húmeda	0	0	933
Costra Salina	0	0	79
Charca	0	0	423
Callaos	0	0	90
Aluviales	0	0	190
Tamarix canar	0	0	226
Traganum moqu	2	0	151
Juncus acutus	0	0	37
Suaeda mollis	3	0	25
Tetraena font	9	0	48
Launaea arbor	0	48	57
Total	14	48	3651

(a)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00
Arena Seca	94.46	0.30	0.00	0.00	0.00	3.19
Arena Húmeda	0.00	94.72	0.00	0.00	0.00	0.00
Costra Salina	2.74	0.00	97.50	0.00	0.00	0.00
Charca	0.00	0.00	0.00	100.00	0.00	0.00
Callaos	0.00	0.00	0.00	0.00	95.74	0.00
Aluviales	2.26	4.87	0.00	0.00	0.00	0.00
Tamarix canar	0.00	0.00	0.00	0.00	0.00	0.00
Traganum moqu	0.00	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	0.00	0.00
Tetraena font	0.55	0.10	0.00	0.00	0.00	1.06
Launaea arbor	0.00	0.00	2.50	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	3.36	0.47	0.59	0.00	0.00
Arena Húmeda	0.00	0.00	0.00	0.00	0.00
Costra Salina	0.00	0.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00
Callaos	0.00	0.00	0.00	0.00	0.00
Aluviales	91.60	0.00	0.00	0.00	0.00
Tamarix canar	0.00	97.64	7.06	15.56	0.00
Traganum moqu	0.00	0.00	87.65	0.00	0.00
Juncus acutus	0.00	0.00	0.00	82.22	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	55.00
Tetraena font	4.20	0.00	3.53	0.00	45.00
Launaea arbor	0.84	1.89	1.18	2.22	0.00
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)			Total
	Tetraena font	Launaea arbor		
Unclassified	0.00	0.00		0.00
Arena Seca	0.00	0.00		38.13
Arena Húmeda	0.00	0.00		25.55
Costra Salina	0.00	0.00		2.16
Charca	0.00	0.00		11.59
Callaos	0.00	0.00		2.47
Aluviales	0.00	0.00		5.20
Tamarix canar	0.00	0.00		6.19
Traganum moqu	14.29	0.00		4.14
Juncus acutus	0.00	0.00		1.01
Suaeda mollis	21.43	0.00		0.68
Tetraena font	64.29	0.00		1.31
Launaea arbor	0.00	100.00		1.56
Total	100.00	100.00		100.00

(b)

Class	Commission		Omission	
	(Percent)	(Pixels)	(Percent)	(Pixels)
Arena Seca	0.86	12/1392	5.54	81/1461
Arena Húmeda	0.00	0/933	5.28	52/985
Costra Salina	50.63	40/79	2.50	1/40
Charca	0.00	0/423	0.00	0/423
Callaos	0.00	0/90	4.26	4/94
Aluviales	42.63	81/190	8.40	10/119
Tamarix canar	8.41	19/226	2.36	5/212
Traganum moqu	1.32	2/151	12.35	21/170
Juncus acutus	0.00	0/37	17.78	8/45
Suaeda mollis	12.00	3/25	45.00	18/40
Tetraena font	81.25	39/48	35.71	5/14
Launaea arbor	15.79	9/57	0.00	0/48

(c)

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Arena Seca	94.46	99.14	1380/1461	1380/1392
Arena Húmeda	94.72	100.00	933/985	933/933
Costra Salina	97.50	49.37	39/40	39/79
Charca	100.00	100.00	423/423	423/423
Callaos	95.74	100.00	90/94	90/90
Aluviales	91.60	57.37	109/119	109/190
Tamarix canar	97.64	91.59	207/212	207/226
Traganum moqu	87.65	98.68	149/170	149/151
Juncus acutus	82.22	100.00	37/45	37/37
Suaeda mollis	55.00	88.00	22/40	22/25
Tetraena font	64.29	18.75	9/14	9/48
Launaea arbor	100.00	84.21	48/48	48/57

(d)

Tabla A2.2. Matriz de Confusión para el clasificador *Maximum Likelihood*, con *Majority*, sobre imagen original: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commision y Omission, (d) Prod. Acc. y User Acc.

Support Vector Machine

Class	Ground Truth (Pixels)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0	0	0	0	0	0
Arena Seca	1431	13	8	0	0	0
Arena Húmeda	8	933	0	0	4	0
Costra Salina	2	0	28	0	0	0
Charca	0	0	0	423	0	0
Callaos	0	3	3	0	88	0
Aluviales	20	36	1	0	2	0
Tamarix canar	0	0	0	0	0	0
Traganum moqu	0	0	0	0	0	0
Juncus acutus	0	0	0	0	0	0
Suaeda mollis	0	0	0	0	0	0
Tetraena font	0	0	0	0	0	0
Launaea arbor	0	0	0	0	0	0
Total	1461	985	40	423	94	

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	23	0	0	0	6
Arena Húmeda	5	0	0	0	0
Costra Salina	0	0	0	0	0
Charca	0	0	0	0	0
Callaos	0	0	0	0	0
Aluviales	91	0	0	0	0
Tamarix canar	0	209	20	6	0
Traganum moqu	0	1	141	0	2
Juncus acutus	0	1	0	39	0
Suaeda mollis	0	1	4	0	13
Tetraena font	0	0	4	0	14
Launaea arbor	0	0	1	0	5
Total	119	212	170	45	40

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	1481
Arena Húmeda	1	0	951
Costra Salina	0	0	30
Charca	0	0	423
Callaos	0	0	94
Aluviales	0	0	150
Tamarix canar	5	0	240
Traganum moqu	2	6	152
Juncus acutus	0	0	40
Suaeda mollis	2	0	20
Tetraena font	3	0	21
Launaea arbor	1	42	49
Total	14	48	3651

(a)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00
Arena Seca	97.95	1.32	20.00	0.00	0.00	0.00
Arena Húmeda	0.55	94.72	0.00	0.00	0.00	4.26
Costra Salina	0.14	0.00	70.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	100.00	0.00	0.00
Callaos	0.00	0.30	7.50	0.00	0.00	93.62
Aluviales	1.37	3.65	2.50	0.00	0.00	2.13
Tamarix canar	0.00	0.00	0.00	0.00	0.00	0.00
Traganum moqu	0.00	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	0.00	0.00
Tetraena font	0.00	0.00	0.00	0.00	0.00	0.00
Launaea arbor	0.00	0.00	0.00	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)					
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis	
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00
Arena Seca	19.33	0.00	0.00	0.00	0.00	15.00
Arena Húmeda	4.20	0.00	0.00	0.00	0.00	0.00
Costra Salina	0.00	0.00	0.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00	0.00
Callaos	0.00	0.00	0.00	0.00	0.00	0.00
Aluviales	76.47	0.00	0.00	0.00	0.00	0.00
Tamarix canar	0.00	98.58	11.76	13.33	0.00	0.00
Traganum moqu	0.00	0.47	82.94	0.00	5.00	0.00
Juncus acutus	0.00	0.47	0.00	86.67	0.00	0.00
Suaeda mollis	0.00	0.47	2.35	0.00	32.50	0.00
Tetraena font	0.00	0.00	2.35	0.00	35.00	0.00
Launaea arbor	0.00	0.00	0.59	0.00	12.50	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)			Total
	Tetraena font	Launaea arbor		
Unclassified	0.00	0.00		0.00
Arena Seca	0.00	0.00		40.56
Arena Húmeda	7.14	0.00		26.05
Costra Salina	0.00	0.00		0.82
Charca	0.00	0.00		11.59
Callaos	0.00	0.00		2.57
Aluviales	0.00	0.00		4.11
Tamarix canar	35.71	0.00		6.57
Traganum moqu	14.29	12.50		4.16
Juncus acutus	0.00	0.00		1.10
Suaeda mollis	14.29	0.00		0.55
Tetraena font	21.43	0.00		0.58
Launaea arbor	7.14	87.50		1.34
Total	100.00	100.00		100.00

(b)

Class	Commission		Omission	
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	3.38	2.05	50/1481	30/1461
Arena Húmeda	1.89	5.28	18/951	52/985
Costra Salina	6.67	30.00	2/30	12/40
Charca	0.00	0.00	0/423	0/423
Callaos	6.38	6.38	6/94	6/94
Aluviales	39.33	23.53	59/150	28/119
Tamarix canar	12.92	1.42	31/240	3/212
Traganum moqu	7.24	17.06	11/152	29/170
Juncus acutus	2.50	13.33	1/40	6/45
Suaeda mollis	35.00	67.50	7/20	27/40
Tetraena font	85.71	78.57	18/21	11/14
Launaea arbor	14.29	12.50	7/49	6/48

(c)

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Arena Seca	97.95	96.62	1431/1461	1431/1481
Arena Húmeda	94.72	98.11	933/985	933/951
Costra Salina	70.00	93.33	28/40	28/30
Charca	100.00	100.00	423/423	423/423
Callaos	93.62	93.62	88/94	88/94
Aluviales	76.47	60.67	91/119	91/150
Tamarix canar	98.58	87.08	209/212	209/240
Tragnum moqu	82.94	92.76	141/170	141/152
Juncus acutus	86.67	97.50	39/45	39/40
Suaeda mollis	32.50	65.00	13/40	13/20
Tetraena font	21.43	14.29	3/14	3/21
Launaea arbor	87.50	85.71	42/48	42/49

(d)

Tabla A2.3. Matriz de Confusión para el clasificador *Support Vector Machine* optimizado sobre imagen original: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commision y Omission, (d) Prod. Acc. y User Acc.

Class	Ground Truth (Pixels)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0	0	0	0	0	0
Arena Seca	1457	3	4	0	0	0
Arena Húmeda	1	953	0	0	0	0
Costra Salina	0	0	35	0	0	0
Charca	0	0	0	423	0	0
Callaos	0	0	1	0	94	0
Aluviales	3	29	0	0	0	0
Tamarix canar	0	0	0	0	0	0
Tragnum moqu	0	0	0	0	0	0
Juncus acutus	0	0	0	0	0	0
Suaeda mollis	0	0	0	0	0	0
Tetraena font	0	0	0	0	0	0
Launaea arbor	0	0	0	0	0	0
Total	1461	985	40	423	94	

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Tragnum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	12	1	1	0	5
Arena Húmeda	1	0	0	0	0
Costra Salina	0	0	0	0	0
Charca	0	0	0	0	0
Callaos	0	0	0	0	0
Aluviales	106	0	0	0	0
Tamarix canar	0	211	10	8	0
Tragnum moqu	0	0	159	0	0
Juncus acutus	0	0	0	37	0
Suaeda mollis	0	0	0	0	9
Tetraena font	0	0	0	0	17
Launaea arbor	0	0	0	0	9
Total	119	212	170	45	40

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	1483
Arena Húmeda	1	0	956
Costra Salina	0	0	35
Charca	0	0	423
Callaos	0	0	95
Aluviales	0	0	138
Tamarix canar	5	0	234
Tragnum moqu	5	0	164
Juncus acutus	0	0	37
Suaeda mollis	1	0	10
Tetraena font	2	0	19
Launaea arbor	0	48	57
Total	14	48	3651

(a)

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)					
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos	
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00
Arena Seca	99.73	0.30	10.00	0.00	0.00	0.00
Arena Húmeda	0.07	96.75	0.00	0.00	0.00	0.00
Costra Salina	0.00	0.00	87.50	0.00	0.00	0.00
Charca	0.00	0.00	0.00	100.00	0.00	0.00
Callaos	0.00	0.00	2.50	0.00	100.00	0.00
Aluviales	0.21	2.94	0.00	0.00	0.00	0.00
Tamarix canar	0.00	0.00	0.00	0.00	0.00	0.00
Traganum moqu	0.00	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	0.00	0.00
Tetraena font	0.00	0.00	0.00	0.00	0.00	0.00
Launaea arbor	0.00	0.00	0.00	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)					
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis	
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00
Arena Seca	10.08	0.47	0.59	0.00	12.50	0.00
Arena Húmeda	0.84	0.00	0.00	0.00	0.00	0.00
Costra Salina	0.00	0.00	0.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00	0.00
Callaos	0.00	0.00	0.00	0.00	0.00	0.00
Aluviales	89.08	0.00	0.00	0.00	0.00	0.00
Tamarix canar	0.00	99.53	5.88	17.78	0.00	0.00
Traganum moqu	0.00	0.00	93.53	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	82.22	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	22.50	0.00
Tetraena font	0.00	0.00	0.00	0.00	42.50	0.00
Launaea arbor	0.00	0.00	0.00	0.00	22.50	0.00
Total	100.00	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)			Total
	Tetraena font	Launaea arbor		
Unclassified	0.00	0.00		0.00
Arena Seca	0.00	0.00		40.62
Arena Húmeda	7.14	0.00		26.18
Costra Salina	0.00	0.00		0.96
Charca	0.00	0.00		11.59
Callaos	0.00	0.00		2.60
Aluviales	0.00	0.00		3.78
Tamarix canar	35.71	0.00		6.41
Traganum moqu	35.71	0.00		4.49
Juncus acutus	0.00	0.00		1.01
Suaeda mollis	7.14	0.00		0.27
Tetraena font	14.29	0.00		0.52
Launaea arbor	0.00	100.00		1.56
Total	100.00	100.00		100.00

(b)

Class	Commission		Omission	
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	1.75	0.27	26/1483	4/1461
Arena Húmeda	0.31	3.25	3/956	32/985
Costra Salina	0.00	12.50	0/35	5/40
Charca	0.00	0.00	0/423	0/423
Callaos	1.05	0.00	1/95	0/94
Aluviales	23.19	10.92	32/138	13/119
Tamarix canar	9.83	0.47	23/234	1/212
Traganum moqu	3.05	6.47	5/164	11/170
Juncus acutus	0.00	17.78	0/37	8/45
Suaeda mollis	10.00	77.50	1/10	31/40
Tetraena font	89.47	85.71	17/19	12/14
Launaea arbor	15.79	0.00	9/57	0/48

(c)

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Arena Seca	99.73	98.25	1457/1461	1457/1483
Arena Húmeda	96.75	99.69	953/985	953/956
Costra Salina	87.50	100.00	35/40	35/35
Charca	100.00	100.00	423/423	423/423
Callaos	100.00	98.95	94/94	94/95
Aluviales	89.08	76.81	106/119	106/138
Tamarix canar	99.53	90.17	211/212	211/234
Traganum moqu	93.53	96.95	159/170	159/164
Juncus acutus	82.22	100.00	37/45	37/37
Suaeda mollis	22.50	90.00	9/40	9/10
Tetraena font	14.29	10.53	2/14	2/19
Launaea arbor	100.00	84.21	48/48	48/57

(d)

Tabla A2.4. Matriz de Confusión para el clasificador *Support Vector Machine* optimizado, con *Majority*, sobre imagen original: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

Linear Spectral Unmixing (LSU)

En este caso, añadiremos una tabla adicional con la leyenda que nos permitirá ver la correspondencia de Class X con las clases de vegetación y sustrato que se han definido en el estudio.

Class	Clase definida	Class	Clase definida
1	Arena seca	7	Aluviales
2	" <i>Launaea arborescens</i> "	8	" <i>Tamarix canariensis</i> "
3	Arena húmeda	9	" <i>Traganum moquinii</i> "
4	Costra Salina	10	" <i>Juncus acutus</i> "
5	Charca	11	" <i>Suaeda mollis</i> "
6	Callaos	12	" <i>Tetraena fontanesii</i> "

Tabla A2.5. Leyenda para la lectura de las matrices de confusión de LSU.

Pasaremos a continuación a exponer las tablas correspondientes a las matrices de confusión obtenidas con la técnica *Linear Spectral Unmixing (LSU)* aplicada sobre la imagen original y posteriormente de manera adicional el filtrado *Majority*.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Pixels)					Charca
	Arena Seca	Launaea arbor	Arena Húmeda	Costra Salina		
Unclassified	0	0	0	0	0	0
Class 1	2550	34	810	32	328	
Class 2	380	24	89	0	0	
Class 3	21	2	211	3	13	
Class 4	60	6	58	28	53	
Class 5	0	3	5	0	141	
Class 6	4	1	67	1	0	
Class 7	0	0	0	0	0	
Class 8	3	4	0	0	0	
Class 9	1	0	0	0	0	
Class 10	0	0	1	0	0	
Class 11	49	13	77	5	0	
Class 12	0	0	0	0	0	
Total	3068	87	1318	69	535	

Class	Ground Truth (Pixels)				
	Callaos	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus
Unclassified	0	0	0	0	0
Class 1	30	39	63	96	30
Class 2	9	10	4	15	3
Class 3	57	3	7	14	3
Class 4	8	5	19	10	4
Class 5	0	0	9	0	2
Class 6	18	3	5	10	0
Class 7	0	0	0	1	0
Class 8	0	0	35	18	2
Class 9	0	0	0	1	0
Class 10	0	0	2	4	9
Class 11	2	8	18	13	9
Class 12	0	0	0	0	0
Total	124	68	162	182	62

Class	Ground Truth (Pixels)			Total
	Suaeda mollis	Tetraena font		
Unclassified	0	0	0	0
Class 1	11	22	4045	
Class 2	0	0	534	
Class 3	1	1	336	
Class 4	3	0	254	
Class 5	2	0	162	
Class 6	0	0	109	
Class 7	0	0	1	
Class 8	0	1	63	
Class 9	0	0	2	
Class 10	0	0	16	
Class 11	8	3	205	
Class 12	0	0	0	
Total	25	27	5727	

(a)

Class	Ground Truth (Percent)					Charca
	Arena Seca	Launaea arbor	Arena Húmeda	Costra Salina		
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00
Class 1	83.12	39.08	61.46	46.38	61.31	
Class 2	12.39	27.59	6.75	0.00	0.00	
Class 3	0.68	2.30	16.01	4.35	2.43	
Class 4	1.96	6.90	4.40	40.58	9.91	
Class 5	0.00	3.45	0.38	0.00	26.36	
Class 6	0.13	1.15	5.08	1.45	0.00	
Class 7	0.00	0.00	0.00	0.00	0.00	
Class 8	0.10	4.60	0.00	0.00	0.00	
Class 9	0.03	0.00	0.00	0.00	0.00	
Class 10	0.00	0.00	0.08	0.00	0.00	
Class 11	1.60	14.94	5.84	7.25	0.00	
Class 12	0.00	0.00	0.00	0.00	0.00	
Total	100.00	100.00	100.00	100.00	100.00	

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)				
	Callaos	Aluviales	Tamarix	canarTraganum	moquJuncus acutus
Unclassified	0.00	0.00	0.00	0.00	0.00
Class 1	24.19	57.35	38.89	52.75	48.39
Class 2	7.26	14.71	2.47	8.24	4.84
Class 3	45.97	4.41	4.32	7.69	4.84
Class 4	6.45	7.35	11.73	5.49	6.45
Class 5	0.00	0.00	5.56	0.00	3.23
Class 6	14.52	4.41	3.09	5.49	0.00
Class 7	0.00	0.00	0.00	0.55	0.00
Class 8	0.00	0.00	21.60	9.89	3.23
Class 9	0.00	0.00	0.00	0.55	0.00
Class 10	0.00	0.00	1.23	2.20	14.52
Class 11	1.61	11.76	11.11	7.14	14.52
Class 12	0.00	0.00	0.00	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)		Total
	Suaeda mollis	Tetraena font	
Unclassified	0.00	0.00	0.00
Class 1	44.00	81.48	70.63
Class 2	0.00	0.00	9.32
Class 3	4.00	3.70	5.87
Class 4	12.00	0.00	4.44
Class 5	8.00	0.00	2.83
Class 6	0.00	0.00	1.90
Class 7	0.00	0.00	0.02
Class 8	0.00	3.70	1.10
Class 9	0.00	0.00	0.03
Class 10	0.00	0.00	0.28
Class 11	32.00	11.11	3.58
Class 12	0.00	0.00	0.00
Total	100.00	100.00	100.00

(b)

Class	Commission (Percent)	Omission (Percent)	Commission (Pixels)	Omission (Pixels)
Class 1	36.96	16.88	1495/4045	518/3068
Class 2	95.51	72.41	510/534	63/87
Class 3	37.20	83.99	125/336	1107/1318
Class 4	88.98	59.42	226/254	41/69
Class 5	12.96	73.64	21/162	394/535
Class 6	83.49	85.48	91/109	106/124
Class 7	100.00	100.00	1/1	68/68
Class 8	44.44	78.40	28/63	127/162
Class 9	50.00	99.45	1/2	181/182
Class 10	43.75	85.48	7/16	53/62
Class 11	96.10	68.00	197/205	17/25
Class 12	0.00	100.00	0/0	27/27

(c)

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Class 1	83.12	63.04	2550/3068	2550/4045
Class 2	27.59	4.49	24/87	24/534
Class 3	16.01	62.80	211/1318	211/336
Class 4	40.58	11.02	28/69	28/254
Class 5	26.36	87.04	141/535	141/162
Class 6	14.52	16.51	18/124	18/109
Class 7	0.00	0.00	0/68	0/1
Class 8	21.60	55.56	35/162	35/63
Class 9	0.55	50.00	1/182	1/2
Class 10	14.52	56.25	9/62	9/16
Class 11	32.00	3.90	8/25	8/205
Class 12	0.00	0.00	0/27	0/0

(d)

Tabla A2.6. Matriz de Confusión para la técnica *Linear Spectral Unmixing* sobre imagen original: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Pixels)		Charca	Callaos	Aluviales
	Arena Seca	Arena Húmeda			
Unclassified	0	0	0	0	0
Class 1	1444	883	376	30	111
Class 3	0	64	0	46	0
Class 5	0	0	46	0	0
Class 6	0	7	0	15	0
Class 7	0	0	0	0	0
Class 8	0	0	0	0	0
Class 9	0	0	0	0	0
Class 10	0	0	0	0	0
Class 11	0	15	0	0	6
Class 12	0	0	0	0	0
Class 2	17	16	0	1	1
Class 4	0	0	1	2	1
Total	1461	985	423	94	119

Class	Ground Truth (Pixels)				
	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis	Tetraena font
Unclassified	0	0	0	0	0
Class 1	126	140	39	38	8
Class 3	0	5	1	0	0
Class 5	1	0	0	0	0
Class 6	0	5	0	0	0
Class 7	0	0	0	0	0
Class 8	31	6	0	0	0
Class 9	0	0	0	0	0
Class 10	0	0	1	0	0
Class 11	45	2	2	2	5
Class 12	0	0	0	0	0
Class 2	3	12	0	0	0
Class 4	6	0	2	0	1
Total	212	170	45	40	14

Class	Ground Truth (Pixels)		Total
	Launaea arbor	Costra Salina	
Unclassified	0	0	0
Class 1	36	29	3260
Class 3	0	1	117
Class 5	0	0	47
Class 6	0	0	27
Class 7	0	0	0
Class 8	0	0	37
Class 9	0	0	0
Class 10	0	0	1
Class 11	7	1	85
Class 12	0	0	0
Class 2	4	0	54
Class 4	1	9	23
Total	48	40	3651

(a)

Class	Ground Truth (Percent)		Charca	Callaos	Aluviales
	Arena Seca	Arena Húmeda			
Unclassified	0.00	0.00	0.00	0.00	0.00
Class 1	98.84	89.64	88.89	31.91	93.28
Class 3	0.00	6.50	0.00	48.94	0.00
Class 5	0.00	0.00	10.87	0.00	0.00
Class 6	0.00	0.71	0.00	15.96	0.00
Class 7	0.00	0.00	0.00	0.00	0.00
Class 8	0.00	0.00	0.00	0.00	0.00
Class 9	0.00	0.00	0.00	0.00	0.00
Class 10	0.00	0.00	0.00	0.00	0.00
Class 11	0.00	1.52	0.00	0.00	5.04
Class 12	0.00	0.00	0.00	0.00	0.00
Class 2	1.16	1.62	0.00	1.06	0.84
Class 4	0.00	0.00	0.24	2.13	0.84
Total	100.00	100.00	100.00	100.00	100.00

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)						
	Tamarix	canar	Traganum	moqu	Juncus acutus	Suaeda mollis	Tetraena font
Unclassified	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Class 1	59.43	82.35	86.67	95.00	57.14		
Class 3	0.00	2.94	2.22	0.00	0.00	0.00	0.00
Class 5	0.47	0.00	0.00	0.00	0.00	0.00	0.00
Class 6	0.00	2.94	0.00	0.00	0.00	0.00	0.00
Class 7	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Class 8	14.62	3.53	0.00	0.00	0.00	0.00	0.00
Class 9	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Class 10	0.00	0.00	2.22	0.00	0.00	0.00	0.00
Class 11	21.23	1.18	4.44	5.00	35.71		
Class 12	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Class 2	1.42	7.06	0.00	0.00	0.00	0.00	0.00
Class 4	2.83	0.00	4.44	0.00	7.14		
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)			Total
	Launaea	arbor	Costra Salina	
Unclassified	0.00	0.00	0.00	0.00
Class 1	75.00	72.50	89.29	
Class 3	0.00	2.50	3.20	
Class 5	0.00	0.00	1.29	
Class 6	0.00	0.00	0.74	
Class 7	0.00	0.00	0.00	
Class 8	0.00	0.00	1.01	
Class 9	0.00	0.00	0.00	
Class 10	0.00	0.00	0.03	
Class 11	14.58	2.50	2.33	
Class 12	0.00	0.00	0.00	
Class 2	8.33	0.00	1.48	
Class 4	2.08	22.50	0.63	
Total	100.00	100.00	100.00	

(b)

Class	Commission (Percent)	Omission (Percent)	Commission (Pixels)	Omission (Pixels)
Class 1	55.71	1.16	1816/3260	17/1461
Class 3	45.30	93.50	53/117	921/985
Class 5	2.13	89.13	1/47	377/423
Class 6	44.44	84.04	12/27	79/94
Class 7	0.00	100.00	0/0	119/119
Class 8	16.22	85.38	6/37	181/212
Class 9	0.00	100.00	0/0	170/170
Class 10	0.00	97.78	0/1	44/45
Class 11	97.65	95.00	83/85	38/40
Class 12	0.00	100.00	0/0	14/14
Class 2	92.59	91.67	50/54	44/48
Class 4	60.87	77.50	14/23	31/40

(c)

Class	Prod. Acc. (Percent)	User Acc. (Percent)	Prod. Acc. (Pixels)	User Acc. (Pixels)
Class 1	98.84	44.29	1444/1461	1444/3260
Class 3	6.50	54.70	64/985	64/117
Class 5	10.87	97.87	46/423	46/47
Class 6	15.96	55.56	15/94	15/27
Class 7	0.00	0.00	0/119	0/0
Class 8	14.62	83.78	31/212	31/37
Class 9	0.00	0.00	0/170	0/0
Class 10	2.22	100.00	1/45	1/1
Class 11	5.00	2.35	2/40	2/85
Class 12	0.00	0.00	0/14	0/0
Class 2	8.33	7.41	4/48	4/54
Class 4	22.50	39.13	9/40	9/23

(d)

Tabla A2.7. Matriz de Confusión para la técnica *Linear Spectral Unmixing*, con *Majority*, sobre imagen original: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

- **Imagen Fusionada**

Maximum Likelihood

Class	Ground Truth (Pixels)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0	0	0	0	0
Arena Seca	21937	118	1	0	74
Arena Húmeda	5	14266	0	0	28
Costra Salina	556	23	286	0	65
Charca	0	0	0	6766	0
Callaos	0	40	0	0	1293
Aluviales	690	592	0	0	0
Tamarix canar	0	0	0	0	14
Traganum moqu	0	0	0	0	0
Juncus acutus	0	0	0	0	0
Suaeda mollis	0	5	0	0	0
Tetraena font	150	715	0	0	30
Launaea arbor	38	1	23	2	0
Total	23376	15760	310	6768	1504

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	366	0	0	0	8
Arena Húmeda	124	0	0	0	0
Costra Salina	2	0	11	0	0
Charca	0	0	0	0	0
Callaos	4	0	0	0	0
Aluviales	1279	0	0	0	0
Tamarix canar	0	3213	256	95	0
Traganum moqu	0	28	2134	0	0
Juncus acutus	0	0	2	586	0
Suaeda mollis	0	0	1	0	351
Tetraena font	129	7	104	17	213
Launaea arbor	0	144	212	22	68
Total	1904	3392	2720	720	640

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	22504
Arena Húmeda	0	0	14423
Costra Salina	1	0	944
Charca	0	0	6766
Callaos	0	0	1337
Aluviales	0	0	2561
Tamarix canar	6	22	3606
Traganum moqu	1	38	2201
Juncus acutus	1	0	589
Suaeda mollis	93	9	459
Tetraena font	102	11	1478
Launaea arbor	20	688	1218
Total	224	768	58086

(a)

Class	Ground Truth (Percent)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	93.84	0.75	0.32	0.00	4.92
Arena Húmeda	0.02	90.52	0.00	0.00	1.86
Costra Salina	2.38	0.15	92.26	0.00	4.32
Charca	0.00	0.00	0.00	99.97	0.00
Callaos	0.00	0.25	0.00	0.00	85.97
Aluviales	2.95	3.76	0.00	0.00	0.00
Tamarix canar	0.00	0.00	0.00	0.00	0.93
Traganum moqu	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.03	0.00	0.00	0.00
Tetraena font	0.64	4.54	0.00	0.00	1.99
Launaea arbor	0.16	0.01	7.42	0.03	0.00
Total	100.00	100.00	100.00	100.00	100.00

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	19.22	0.00	0.00	0.00	1.25
Arena Húmeda	6.51	0.00	0.00	0.00	0.00
Costra Salina	0.11	0.00	0.40	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00
Callaos	0.21	0.00	0.00	0.00	0.00
Aluviales	67.17	0.00	0.00	0.00	0.00
Tamarix canar	0.00	94.72	9.41	13.19	0.00
Traganum moqu	0.00	0.83	78.46	0.00	0.00
Juncus acutus	0.00	0.00	0.07	81.39	0.00
Suaeda mollis	0.00	0.00	0.04	0.00	54.84
Tetraena font	6.78	0.21	3.82	2.36	33.28
Launaea arbor	0.00	4.25	7.79	3.06	10.63
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)		Total
	Tetraena font	Launaea arbor	
Unclassified	0.00	0.00	0.00
Arena Seca	0.00	0.00	38.74
Arena Húmeda	0.00	0.00	24.83
Costra Salina	0.45	0.00	1.63
Charca	0.00	0.00	11.65
Callaos	0.00	0.00	2.30
Aluviales	0.00	0.00	4.41
Tamarix canar	2.68	2.86	6.21
Traganum moqu	0.45	4.95	3.79
Juncus acutus	0.45	0.00	1.01
Suaeda mollis	41.52	1.17	0.79
Tetraena font	45.54	1.43	2.54
Launaea arbor	8.93	89.58	2.10
Total	100.00	100.00	100.00

(b)

Class	Commission		Omission	
	(Percent)	(Pixels)	(Percent)	(Pixels)
Arena Seca	2.52	567/22504	6.16	1439/23376
Arena Húmeda	1.09	157/14423	9.48	1494/15760
Costra Salina	69.70	658/944	7.74	24/310
Charca	0.00	0/6766	0.03	2/6768
Callaos	3.29	44/1337	14.03	211/1504
Aluviales	50.06	1282/2561	32.83	625/1904
Tamarix canar	10.90	393/3606	5.28	179/3392
Traganum moqu	3.04	67/2201	21.54	586/2720
Juncus acutus	0.51	3/589	18.61	134/720
Suaeda mollis	23.53	108/459	45.16	289/640
Tetraena font	93.10	1376/1478	54.46	122/224
Launaea arbor	43.51	530/1218	10.42	80/768

(c)

Class	Prod. Acc.		User Acc.	
	(Percent)	(Pixels)	(Percent)	(Pixels)
Arena Seca	93.84	21937/23376	97.48	21937/22504
Arena Húmeda	90.52	14266/15760	98.91	14266/14423
Costra Salina	92.26	286/310	30.30	286/944
Charca	99.97	6766/6768	100.00	6766/6766
Callaos	85.97	1293/1504	96.71	1293/1337
Aluviales	67.17	1279/1904	49.94	1279/2561
Tamarix canar	94.72	3213/3392	89.10	3213/3606
Traganum moqu	78.46	2134/2720	96.96	2134/2201
Juncus acutus	81.39	586/720	99.49	586/589
Suaeda mollis	54.84	351/640	76.47	351/459
Tetraena font	45.54	102/224	6.90	102/1478
Launaea arbor	89.58	688/768	56.49	688/1218

(d)

Tabla A2.8. Matriz de Confusión para el clasificador *Maximum Likelihood* sobre imagen fusionada: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Pixels)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0	0	0	0	0
Arena Seca	22026	82	0	0	67
Arena Húmeda	0	14513	0	0	26
Costra Salina	550	12	287	0	57
Charca	0	0	0	6766	0
Callaos	0	22	0	0	1312
Aluviales	691	579	0	0	0
Tamarix canar	0	0	0	0	14
Traganum moqu	0	0	0	0	0
Juncus acutus	0	0	0	0	0
Suaeda mollis	0	4	0	0	0
Tetraena font	95	548	0	0	28
Launaea arbor	14	0	23	2	0
Total	23376	15760	310	6768	1504

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	368	0	1	0	7
Arena Húmeda	110	0	0	0	0
Costra Salina	0	0	10	0	0
Charca	0	0	0	0	0
Callaos	2	0	0	0	0
Aluviales	1298	0	0	0	0
Tamarix canar	0	3224	218	97	0
Traganum moqu	0	18	2219	0	0
Juncus acutus	0	0	0	586	0
Suaeda mollis	0	0	0	0	354
Tetraena font	126	5	78	16	232
Launaea arbor	0	145	194	21	47
Total	1904	3392	2720	720	640

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	22551
Arena Húmeda	0	0	14649
Costra Salina	0	0	916
Charca	0	0	6766
Callaos	0	0	1336
Aluviales	0	0	2568
Tamarix canar	6	13	3572
Traganum moqu	0	28	2265
Juncus acutus	0	0	586
Suaeda mollis	97	1	456
Tetraena font	102	6	1236
Launaea arbor	19	720	1185
Total	224	768	58086

(a)

Class	Ground Truth (Percent)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	94.22	0.52	0.00	0.00	4.45
Arena Húmeda	0.00	92.09	0.00	0.00	1.73
Costra Salina	2.35	0.08	92.58	0.00	3.79
Charca	0.00	0.00	0.00	99.97	0.00
Callaos	0.00	0.14	0.00	0.00	87.23
Aluviales	2.96	3.67	0.00	0.00	0.00
Tamarix canar	0.00	0.00	0.00	0.00	0.93
Traganum moqu	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.03	0.00	0.00	0.00
Tetraena font	0.41	3.48	0.00	0.00	1.86
Launaea arbor	0.06	0.00	7.42	0.03	0.00
Total	100.00	100.00	100.00	100.00	100.00

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	19.33	0.00	0.04	0.00	1.09
Arena Húmeda	5.78	0.00	0.00	0.00	0.00
Costra Salina	0.00	0.00	0.37	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00
Callaos	0.11	0.00	0.00	0.00	0.00
Aluviales	68.17	0.00	0.00	0.00	0.00
Tamarix canar	0.00	95.05	8.01	13.47	0.00
Traganum moqu	0.00	0.53	81.58	0.00	0.00
Juncus acutus	0.00	0.00	0.00	81.39	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	55.31
Tetraena font	6.62	0.15	2.87	2.22	36.25
Launaea arbor	0.00	4.27	7.13	2.92	7.34
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)		Total
	Tetraena font	Launaea arbor	
Unclassified	0.00	0.00	0.00
Arena Seca	0.00	0.00	38.82
Arena Húmeda	0.00	0.00	25.22
Costra Salina	0.00	0.00	1.58
Charca	0.00	0.00	11.65
Callaos	0.00	0.00	2.30
Aluviales	0.00	0.00	4.42
Tamarix canar	2.68	1.69	6.15
Traganum moqu	0.00	3.65	3.90
Juncus acutus	0.00	0.00	1.01
Suaeda mollis	43.30	0.13	0.79
Tetraena font	45.54	0.78	2.13
Launaea arbor	8.48	93.75	2.04
Total	100.00	100.00	100.00

(b)

Class	Commission	Omission	Commission	Omission
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	2.33	5.78	525/22551	1350/23376
Arena Húmeda	0.93	7.91	136/14649	1247/15760
Costra Salina	68.67	7.42	629/916	23/310
Charca	0.00	0.03	0/6766	2/6768
Callaos	1.80	12.77	24/1336	192/1504
Aluviales	49.45	31.83	1270/2568	606/1904
Tamarix canar	9.74	4.95	348/3572	168/3392
Traganum moqu	2.03	18.42	46/2265	501/2720
Juncus acutus	0.00	18.61	0/586	134/720
Suaeda mollis	22.37	44.69	102/456	286/640
Tetraena font	91.75	54.46	1134/1236	122/224
Launaea arbor	39.24	6.25	465/1185	48/768

(c)

Class	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	94.22	97.67	22026/23376	22026/22551
Arena Húmeda	92.09	99.07	14513/15760	14513/14649
Costra Salina	92.58	31.33	287/310	287/916
Charca	99.97	100.00	6766/6768	6766/6766
Callaos	87.23	98.20	1312/1504	1312/1336
Aluviales	68.17	50.55	1298/1904	1298/2568
Tamarix canar	95.05	90.26	3224/3392	3224/3572
Traganum moqu	81.58	97.97	2219/2720	2219/2265
Juncus acutus	81.39	100.00	586/720	586/586
Suaeda mollis	55.31	77.63	354/640	354/456
Tetraena font	45.54	8.25	102/224	102/1236
Launaea arbor	93.75	60.76	720/768	720/1185

(d)

Tabla A2.9. Matriz de Confusión para el clasificador *Maximum Likelihood*, con *Majority*, sobre imagen fusionada: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

Support Vector Machine

Class	Ground Truth (Pixels)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0	0	0	0	0
Arena Seca	22265	468	47	0	1
Arena Húmeda	258	14891	0	0	71
Costra Salina	231	0	239	0	0
Charca	0	0	0	6768	0
Callaos	1	19	23	0	1419
Aluviales	621	316	0	0	13
Tamarix canar	0	0	0	0	0
Traganum moqu	0	0	0	0	0
Juncus acutus	0	0	0	0	0
Suaeda mollis	0	0	0	0	0
Tetraena font	0	66	0	0	0
Launaea arbor	0	0	1	0	0
Total	23376	15760	310	6768	1504

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	594	0	17	0	5
Arena Húmeda	307	0	0	0	5
Costra Salina	0	0	0	0	0
Charca	0	0	0	0	0
Callaos	1	0	0	0	0
Aluviales	1002	0	5	0	0
Tamarix canar	0	3220	378	83	0
Traganum moqu	0	54	2146	0	0
Juncus acutus	0	6	3	637	0
Suaeda mollis	0	6	64	0	282
Tetraena font	0	96	60	0	177
Launaea arbor	0	10	47	0	171
Total	1904	3392	2720	720	640

Class	Ground Truth (Pixels)			Total
	Tetraena font	Launaea arbor		
Unclassified	0	0		0
Arena Seca	0	0		23397
Arena Húmeda	1	0		15533
Costra Salina	0	0		470
Charca	0	0		6768
Callaos	0	0		1463
Aluviales	0	0		1957
Tamarix canar	62	14		3757
Traganum moqu	33	94		2327
Juncus acutus	0	0		646
Suaeda mollis	46	37		435
Tetraena font	43	50		492
Launaea arbor	39	573		841
Total	224	768		58086

(a)

Class	Ground Truth (Percent)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	95.25	2.97	15.16	0.00	0.07
Arena Húmeda	1.10	94.49	0.00	0.00	4.72
Costra Salina	0.99	0.00	77.10	0.00	0.00
Charca	0.00	0.00	0.00	100.00	0.00
Callaos	0.00	0.12	7.42	0.00	94.35
Aluviales	2.66	2.01	0.00	0.00	0.86
Tamarix canar	0.00	0.00	0.00	0.00	0.00
Traganum moqu	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	0.00
Tetraena font	0.00	0.42	0.00	0.00	0.00
Launaea arbor	0.00	0.00	0.32	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	31.20	0.00	0.63	0.00	0.78
Arena Húmeda	16.12	0.00	0.00	0.00	0.78
Costra Salina	0.00	0.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00
Callaos	0.05	0.00	0.00	0.00	0.00
Aluviales	52.63	0.00	0.18	0.00	0.00
Tamarix canar	0.00	94.93	13.90	11.53	0.00
Traganum moqu	0.00	1.59	78.90	0.00	0.00
Juncus acutus	0.00	0.18	0.11	88.47	0.00
Suaeda mollis	0.00	0.18	2.35	0.00	44.06
Tetraena font	0.00	2.83	2.21	0.00	27.66
Launaea arbor	0.00	0.29	1.73	0.00	26.72
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)		Total
	Tetraena font	Launaea arbor	
Unclassified	0.00	0.00	0.00
Arena Seca	0.00	0.00	40.28
Arena Húmeda	0.45	0.00	26.74
Costra Salina	0.00	0.00	0.81
Charca	0.00	0.00	11.65
Callaos	0.00	0.00	2.52
Aluviales	0.00	0.00	3.37
Tamarix canar	27.68	1.82	6.47
Traganum moqu	14.73	12.24	4.01
Juncus acutus	0.00	0.00	1.11
Suaeda mollis	20.54	4.82	0.75
Tetraena font	19.20	6.51	0.85
Launaea arbor	17.41	74.61	1.45
Total	100.00	100.00	100.00

(b)

Class	Commission		Omission	
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	4.84	4.75	1132/23397	1111/23376
Arena Húmeda	4.13	5.51	642/15533	869/15760
Costra Salina	49.15	22.90	231/470	71/310
Charca	0.00	0.00	0/6768	0/6768
Callaos	3.01	5.65	44/1463	85/1504
Aluviales	48.80	47.37	955/1957	902/1904
Tamarix canar	14.29	5.07	537/3757	172/3392
Traganum moqu	7.78	21.10	181/2327	574/2720
Juncus acutus	1.39	11.53	9/646	83/720
Suaeda mollis	35.17	55.94	153/435	358/640
Tetraena font	91.26	80.80	449/492	181/224
Launaea arbor	31.87	25.39	268/841	195/768

(c)

Class	Prod. Acc.		User Acc.	
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	95.25	95.16	22265/23397	22265/23397
Arena Húmeda	94.49	95.87	14891/15760	14891/15533
Costra Salina	77.10	50.85	239/310	239/470
Charca	100.00	100.00	6768/6768	6768/6768
Callaos	94.35	96.99	1419/1504	1419/1463
Aluviales	52.63	51.20	1002/1904	1002/1957
Tamarix canar	94.93	85.71	3220/3392	3220/3757
Traganum moqu	78.90	92.22	2146/2720	2146/2327
Juncus acutus	88.47	98.61	637/720	637/646
Suaeda mollis	44.06	64.83	282/640	282/435
Tetraena font	19.20	8.74	43/224	43/492
Launaea arbor	74.61	68.13	573/768	573/841

(d)

Tabla A2.10. Matriz de Confusión para el clasificador *Support Vector Machine* sobre imagen fusionada: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Pixels)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0	0	0	0	0
Arena Seca	22502	355	38	0	2
Arena Húmeda	138	15129	0	0	54
Costra Salina	228	0	248	0	0
Charca	0	0	0	6768	0
Callaos	0	0	23	0	1447
Aluviales	508	222	0	0	1
Tamarix canar	0	0	0	0	0
Traganum moqu	0	0	0	0	0
Juncus acutus	0	0	0	0	0
Suaeda mollis	0	0	0	0	0
Tetraena font	0	54	0	0	0
Launaea arbor	0	0	1	0	0
Total	23376	15760	310	6768	1504

Class	Ground Truth (Pixels)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0	0	0	0	0
Arena Seca	536	0	26	0	5
Arena Húmeda	242	0	0	0	1
Costra Salina	0	0	0	0	0
Charca	0	0	0	0	0
Callaos	0	0	0	0	0
Aluviales	1126	0	0	0	0
Tamarix canar	0	3238	338	90	0
Traganum moqu	0	43	2267	0	0
Juncus acutus	0	0	0	630	0
Suaeda mollis	0	1	35	0	306
Tetraena font	0	101	22	0	152
Launaea arbor	0	9	32	0	176
Total	1904	3392	2720	720	640

Class	Ground Truth (Pixels)		Total
	Tetraena font	Launaea arbor	
Unclassified	0	0	0
Arena Seca	0	0	23464
Arena Húmeda	0	0	15564
Costra Salina	0	0	476
Charca	0	0	6768
Callaos	0	0	1470
Aluviales	0	0	1857
Tamarix canar	67	12	3745
Traganum moqu	32	73	2415
Juncus acutus	0	0	630
Suaeda mollis	50	28	420
Tetraena font	38	36	403
Launaea arbor	37	619	874
Total	224	768	58086

(a)

Class	Ground Truth (Percent)				
	Arena Seca	Arena Húmeda	Costra Salina	Charca	Callaos
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	96.26	2.25	12.26	0.00	0.13
Arena Húmeda	0.59	96.00	0.00	0.00	3.59
Costra Salina	0.98	0.00	80.00	0.00	0.00
Charca	0.00	0.00	0.00	100.00	0.00
Callaos	0.00	0.00	7.42	0.00	96.21
Aluviales	2.17	1.41	0.00	0.00	0.07
Tamarix canar	0.00	0.00	0.00	0.00	0.00
Traganum moqu	0.00	0.00	0.00	0.00	0.00
Juncus acutus	0.00	0.00	0.00	0.00	0.00
Suaeda mollis	0.00	0.00	0.00	0.00	0.00
Tetraena font	0.00	0.34	0.00	0.00	0.00
Launaea arbor	0.00	0.00	0.32	0.00	0.00
Total	100.00	100.00	100.00	100.00	100.00

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

Class	Ground Truth (Percent)				
	Aluviales	Tamarix canar	Traganum moqu	Juncus acutus	Suaeda mollis
Unclassified	0.00	0.00	0.00	0.00	0.00
Arena Seca	28.15	0.00	0.96	0.00	0.78
Arena Húmeda	12.71	0.00	0.00	0.00	0.16
Costra Salina	0.00	0.00	0.00	0.00	0.00
Charca	0.00	0.00	0.00	0.00	0.00
Callaos	0.00	0.00	0.00	0.00	0.00
Aluviales	59.14	0.00	0.00	0.00	0.00
Tamarix canar	0.00	95.46	12.43	12.50	0.00
Traganum moqu	0.00	1.27	83.35	0.00	0.00
Juncus acutus	0.00	0.00	0.00	87.50	0.00
Suaeda mollis	0.00	0.03	1.29	0.00	47.81
Tetraena font	0.00	2.98	0.81	0.00	23.75
Launaea arbor	0.00	0.27	1.18	0.00	27.50
Total	100.00	100.00	100.00	100.00	100.00

Class	Ground Truth (Percent)		Total
	Tetraena font	Launaea arbor	
Unclassified	0.00	0.00	0.00
Arena Seca	0.00	0.00	40.40
Arena Húmeda	0.00	0.00	26.79
Costra Salina	0.00	0.00	0.82
Charca	0.00	0.00	11.65
Callaos	0.00	0.00	2.53
Aluviales	0.00	0.00	3.20
Tamarix canar	29.91	1.56	6.45
Traganum moqu	14.29	9.51	4.16
Juncus acutus	0.00	0.00	1.08
Suaeda mollis	22.32	3.65	0.72
Tetraena font	16.96	4.69	0.69
Launaea arbor	16.52	80.60	1.50
Total	100.00	100.00	100.00

(b)

Class	Commission	Omission	Commission	Omission
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	4.10	3.74	962/23464	874/23376
Arena Húmeda	2.79	4.00	435/15564	631/15760
Costra Salina	47.90	20.00	228/476	62/310
Charca	0.00	0.00	0/6768	0/6768
Callaos	1.56	3.79	23/1470	57/1504
Aluviales	39.36	40.86	731/1857	778/1904
Tamarix canar	13.54	4.54	507/3745	154/3392
Traganum moqu	6.13	16.65	148/2415	453/2720
Juncus acutus	0.00	12.50	0/630	90/720
Suaeda mollis	27.14	52.19	114/420	334/640
Tetraena font	90.57	83.04	365/403	186/224
Launaea arbor	29.18	19.40	255/874	149/768

(c)

Class	Prod. Acc.	User Acc.	Prod. Acc.	User Acc.
	(Percent)	(Percent)	(Pixels)	(Pixels)
Arena Seca	96.26	95.90	22502/23376	22502/23464
Arena Húmeda	96.00	97.21	15129/15760	15129/15564
Costra Salina	80.00	52.10	248/310	248/476
Charca	100.00	100.00	6768/6768	6768/6768
Callaos	96.21	98.44	1447/1504	1447/1470
Aluviales	59.14	60.64	1126/1904	1126/1857
Tamarix canar	95.46	86.46	3238/3392	3238/3745
Traganum moqu	83.35	93.87	2267/2720	2267/2415
Juncus acutus	87.50	100.00	630/720	630/630
Suaeda mollis	47.81	72.86	306/640	306/420
Tetraena font	16.96	9.43	38/224	38/403
Launaea arbor	80.60	70.82	619/768	619/874

(d)

Tabla A2.11. Matriz de Confusión para el clasificador *Support Vector Machine*, con *Majority*, sobre imagen fusionada: (a) Ground Truth (Pixels), (b) Ground Truth (Percent), (c) Commission y Omission, (d) Prod. Acc. y User Acc.

Anexo 3

Código de Matlab

Durante la memoria, hemos explicado que ha sido necesaria la generación de un código en la herramienta software Matlab. Este código ha sido creado con el objetivo de tener las cartografías resultantes en una gama de colores lógicos y razonables de cara a la clase que representan. Por tanto, la funcionalidad del mismo es bastante sencilla.

En primer lugar lee la imagen de interés en el directorio correspondiente. Define los colores que debemos buscar en la imagen que nos ha proporcionado ENVI y le hace una modificación sencilla a dicha imagen para facilitar su lectura. A continuación, recorre píxel a píxel la imagen calculando la diferencia de color entre ese píxel y las 12 clases definidas. El color respecto al cual su diferencia sea menor nos dirá la clase a la que corresponde. Una vez sepamos la clase a la que corresponde el píxel en estudio, se le asigna el nuevo color deseado. Este proceso se repite hasta que recorremos todos los píxeles de nuestra imagen. Realizando un nuevo reshape, tendríamos la imagen final.

Código

```
f=imread('c:/TRABAJO FIN DE MASTER\Resultados ENVI\Imagen Junio 2015\JPEG\
SVM_Cygammapordefecto_Mascara_y_Majority.jpg');
copiaimagen=f;
[Filas,Columnas,RGB]=size(copiaimagen);
imfinal=zeros(Filas,Columnas,RGB);

%COLORES QUE BUSCAMOS EN LA IMAGEN CLASIFICADA
n=[0 0 0]; %Negro de la máscara

%Sustrato de las Dunas
as=[255 255 0]; %Arena Seca
ah=[216 106 53]; %Arena Húmeda
cs=[255 255 230]; %Costra Salina
ch=[0 204 211]; %Charca
ca=[136 141 0]; %Callaos
al=[44 145 89]; %Aluviales

%Vegetación de las Dunas
tc=[0 163 0]; %Tamarix Canariensis
tm=[0 140 125]; %Traganum Moquinii
ja=[13 243 233]; %Juncus Acutus
sm=[180 110 208]; %Suaeda Mollis
tf=[149 255 194]; %Tetraena Fontanesii
la=[203 207 0]; %Launaea Arborescens

matrix_image=reshape(copiaimagen,Filas*Columnas,RGB);

for i = 1 : 1 : (Filas*Columnas)
    acumulador=zeros(1,13);

    for k = 1 : 1 : 3

        acumulador (1)=acumulador(1)+abs(double(matrix_image(i,k))-as(k)); %Arena
Seca
        acumulador (2)=acumulador(2)+abs(double(matrix_image(i,k))-ah(k)); %Arena
Húmeda
        acumulador (3)=acumulador(3)+abs(double(matrix_image(i,k))-cs(k));
%Costra Sali
        acumulador (4)=acumulador(4)+abs(double(matrix_image(i,k))-ch(k));
%Charca
```

Evaluación de técnicas avanzadas de clasificación para la generación de cartografía de alta resolución.

```
        acumulador (5)=acumulador(5)+abs(double(matrix_image(i,k))-ca(k));
%Callaos
        acumulador (6)=acumulador(6)+abs(double(matrix_image(i,k))-al(k));
%Aluviales
        acumulador (7)=acumulador(7)+abs(double(matrix_image(i,k))-tc(k));
%Tamarix Ca
        acumulador (8)=acumulador(8)+abs(double(matrix_image(i,k))-tm(k));
%Traganum Mo
        acumulador (9)=acumulador(9)+abs(double(matrix_image(i,k))-ja(k));
%Juncus Acut
        acumulador (10)=acumulador(10)+abs(double(matrix_image(i,k))-sm(k));
%Suaeda Mo
        acumulador (11)=acumulador(11)+abs(double(matrix_image(i,k))-tf(k));
%Tetraena
        acumulador (12)=acumulador(12)+abs(double(matrix_image(i,k))-la(k));
%Launaea A
        acumulador (13)=acumulador(13)+abs(double(matrix_image(i,k))-n(k));
%Negro másc

        [~, posicion]=min(acumulador);
    end

    switch posicion
        case 1
            matrix_image(i,:)=[240 240 140]; %Nuevo color Arena Seca
        case 2
            matrix_image(i,:)=[205 205 177]; %Nuevo color Arena Húmeda
        case 3
            matrix_image(i,:)=[255 255 255]; %Nuevo color Costra Salina
        case 4
            matrix_image(i,:)=[0 100 200]; %Nuevo color Charca
        case 5
            matrix_image(i,:)=[145 145 145]; %Nuevo color Callaos
        case 6
            matrix_image(i,:)=[196 127 0]; %Nuevo color Aluviales
        case 7
            matrix_image(i,:)=[0 128 0]; %Nuevo color Tamarix Canariensis
        case 8
            matrix_image(i,:)=[110 190 160]; %Nuevo color Traganum Moquinii
        case 9
            matrix_image(i,:)=[0 220 200]; %Nuevo color Juncus Acutus
        case 10
            matrix_image(i,:)=[128 128 192]; %Nuevo color Suaeda Mollis
        case 11
            matrix_image(i,:)=[180 200 250]; %Nuevo color Tetraena Fontanesii
        case 12
            matrix_image(i,:)=[196 196 0]; %Nuevo color Launaea Arborescens
        case 13
            matrix_image(i,:)=[0 0 0]; %Mantenemos el negro
        otherwise
            matrix_image(i,:)=[255 0 0]; % ROJO
    end
end

imfinal=reshape(matrix_image,Filas,Columnas,RGB);
```

