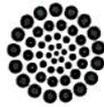




GOBIERNO DE
MÉXICO



CONAHCYT
CONSEJO NACIONAL DE HUMANIDADES
CIENCIAS Y TECNOLOGÍAS

INFOTEC

AUTORIZACIÓN DE IMPRESIÓN Y NO ADEUDO EN BIBLIOTECA

Doctorado en Ciencias en Ciencia de Datos (DCCD)

Ciudad de México, 11 de mayo de 2023

Unidad de Posgrados
PRESENTE

Por medio de la presente se hace constar que el trabajo de titulación

“Evaluación de características estadísticas, de textura y de forma para la clasificación del uso y cobertura de suelo sobre imágenes Landsat y Sentinel-2”

Desarrollado por el alumno: Abel Alejandro Coronado Iruegas, y bajo la asesoría de la Dra. Daniela Alejandra Moctezuma Ochoa cumple con el formato de Biblioteca. Por lo cual, se expide la presente autorización para impresión del proyecto terminal al que se ha hecho mención. Asimismo, se hace constar que no adeuda materiales de la biblioteca de INFOTEC.

No omito mencionar, que se deberá anexar la presente autorización al inicio de la versión impresa del trabajo referido, con el fin de amparar la misma.

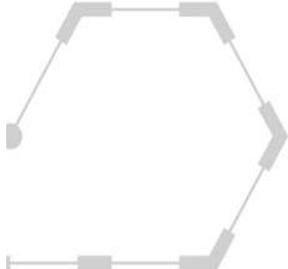
Sin más por el momento, aprovecho la ocasión para enviar un cordial saludo.

Mtro. Carlos Josué Lavandeira Portillo
Director Adjunto de Innovación y Conocimiento

CJLP/jah

C.c.p. Felipe Alfonso Delgado Castillo.- Gerente de Capital Humano.- Para su conocimiento
Abel Alejandro Coronado Iruegas.- Alumno del Doctorado en Ciencias en Ciencia de Datos.- Para su conocimiento





INFOTEC CENTRO DE INVESTIGACIÓN E
INNOVACIÓN EN TECNOLOGÍAS DE LA
INFORMACIÓN Y COMUNICACIÓN

DIRECCIÓN ADJUNTA DE INNOVACIÓN Y
CONOCIMIENTO
GERENCIA DE CAPITAL HUMANO
POSGRADOS

**Evaluación de
características
estadísticas, de textura
y de forma para la
clasificación del uso y
cobertura de suelo
sobre imágenes
Landsat y Sentinel-2**

Tesis

Que para obtener el grado de DOCTOR EN
CIENCIAS EN CIENCIA DE DATOS

Presenta:

Abel Alejandro Coronado Iruegas

Asesor:

Dra. Daniela Alejandra Moctezuma Ochoa

Aguascalientes, abril 2023

Resumen

El uso de datos de percepción remota se ha vuelto muy útil para generar información estadística sobre la sociedad y su entorno. En este sentido, el uso del suelo y la clasificación de la cobertura del suelo son tareas relacionadas con la determinación de la cubierta en la superficie de la Tierra. En el proceso de toma de decisiones, este tipo de información es relevante para mejorar el aprovechamiento de la información sobre eventos como desastres naturales o información catastral, etc.

Para asistir a esto, en esta tesis se propone una metodología para realizar una clasificación supervisada basada en la combinación de características estadísticas, de textura y de forma, para el problema de clasificación de uso y cobertura de suelo (LULC, por sus siglas en inglés).

Para los experimentos, se utilizaron miles de imágenes Landsat y Sentinel-2 que cubren a México y Europa, respectivamente. Para el territorio de México se establecieron 12 clases LULC y 10 para Europa. Utilizando el mismo enfoque de extracción de características y una evaluación con múltiples clasificadores (Random Forest, Support Vector Machines, Extra-Trees, y Artificial Neural Networks) para ambos conjuntos de datos, se logró un 77.1 % de exactitud general para los experimentos Landsat y un 96.7 % para Sentinel-2.

El rendimiento de los diferentes tipos de características se comparó utilizando imágenes de México y Europa. En los resultados se midió la relevancia de cada tipo de característica en función de su participación en el conjunto que logró el mejor desempeño.

Como resultado de las actividades del doctorado se publicaron dos trabajos, un trabajo arbitrado y una base de datos con imágenes Landsat etiquetadas, adicionalmente se generó una imagen nacional libre de nubes:

1. Abel Coronado and Daniela Moctezuma “Feature evaluation for land use and land cover classification based on statistical, textural, and shape features over

Landsat and Sentinel imagery” Journal of Applied Remote Sensing 14(4), 048503 (16 November 2020). <https://doi.org/10.1117/1.JRS.14.048503>

2. Abel Coronado and Daniela Moctezuma “Multispectral and augmented Landsat data with land cover labels (1.0) [Data set]” Zenodo (12 June 2020). <https://doi.org/10.5281/ZENODO.3891579>
3. Abel Coronado and Daniela Moctezuma “Cloud free multispectral Landsat geomedian of Mexico for the year 2011 [Data set]”, Zenodo, (16 October 2022). <https://doi.org/10.5281/zenodo.7213479>

Agradecimientos

Agradezco a mi familia, Rocío y Adriana, por todo su amor y ser la fuente de mi inspiración y fortaleza.

A mis padres, Abel y Virginia, por siempre creer en mí y apoyarme en todo momento.

También agradezco a mi directora de tesis Daniela por su paciencia, perseverancia y acertados consejos.

A mis compañeros de trabajo y en especial, a Juan, Silvia, Bucio, Elio, Gerardo, Sergio, Jimena, Paloma y Julio por apoyar incondicionalmente el trabajo realizado en esta Tesis.

Agradezco la paciencia y generosidad de mis profesores en el Doctorado: Mario, Eric, Sabino, Daniel y Magali.

Y por último a mis compañeros de estudios con quienes compartimos 4 años de emociones y trabajo duro: Claudia, José, Vladimir y todos con quienes compartí esta maravillosa aventura.

Agradezco el apoyo otorgado mediante el proyecto FORDECYT 296737 “Consortio en Inteligencia Artificial” para el desarrollo de la presente tesis.

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Descripción del Problema	4
1.3. Objetivos	6
1.4. Contribuciones	6
1.5. Publicaciones	8
1.6. Esquema de la Tesis	8
2. Fundamentos	11
2.1. Percepción Remota	11
2.1.1. Características de los satélites y sus sensores	13
2.1.2. Características de las observaciones de percepción remota	14
2.1.3. Niveles de procesamiento de los datos satelitales	15
2.2. Fundamentos de aprendizaje automático	16
2.2.1. Algoritmo de clasificación supervisada	17
2.3. Selección de Características	29
2.3.1. Selección de características para dos clases	29
2.3.2. Selección de características multiclase	31
2.3.3. Selección de características para muy altas dimensiones	32
2.4. Resumen	34
3. Trabajos Relacionados	39
3.1. Problemas de aprendizaje Supervisado en Percepción Remota	39
3.2. Resumen	51
4. Bases de Datos de Imágenes Satelitales Landsat para México y Sentinel-2 para Europa	56
4.1. Imágenes Landsat	56
4.1.1. Procedimiento de construcción del conjunto de datos	58

4.1.2. Cubo de Datos Geoespaciales de Imágenes de Satélite	67
4.1.3. Publicación de la Base de Datos	70
4.2. Imágenes Sentinel-2	71
4.3. Resumen	73
5. Clasificación de cobertura y uso de suelo	75
5.1. Extracción de características	76
5.2. Aumento espectral	77
5.3. Características basadas en resúmenes estadísticos	79
5.4. Características basadas en textura	80
5.4.1. Matriz de co-ocurrencia del nivel de gris (GLCM, por sus siglas en inglés de <i>Gray-Level Co-occurrence Matrix</i>)	81
5.4.2. Patrones binarios locales (LBP, por sus siglas en inglés de <i>Local Bi- nary Patterns</i>) + GLCM	83
5.4.3. Características usando filtros de Gabor	84
5.5. Características de forma	85
5.5.1. Momentos de Hu	85
5.6. Selección de Características y Proceso de Clasificación	87
5.7. Resumen	88
6. Experimentos y Resultados	91
6.1. Experimentos	91
6.2. Resultados	92
6.3. Reporte de variación en los experimentos	99
6.3.1. Variación en la exactitud	99
6.3.2. Selección de características	100
6.3.3. Variabilidad de las muestras de imágenes	102
6.4. Resumen	104
Conclusiones y trabajo futuro	108
Bibliografía	122

Índice de figuras

2.1. Comparación de la longitud de onda, frecuencia y energía del espectro electromagnético (crédito: imagina el universo de la NASA).	11
2.2. Ejemplos de firmas espectrales. Las franjas corresponden a las bandas espectrales Landsat: 1 azul, 2 verde, 3 rojo, 4 infrarojo, 5 SWIR 1 y 7 SWIR 2 (Crédito: ARSET de la NASA).	12
2.3. Comparación de la resolución espacial de los satélites Landsat y Sentinel-2, (elaboración propia).	14
2.4. Ejemplo de segmentación del espacio de características usando decisiones binarias; 1) Segmentación del espacio de características a partir de múltiples umbrales, 2) Árbol de decisión correspondiente, el cual muestra la secuencia de preguntas binarias sobre el espacio de características, (elaboración propia).	18
2.5. Ejemplo de segmentación del espacio de características límites de decisión: 1) del lado izquierdo se muestran múltiples fronteras de decisión donde se puede ver que están muy pegadas a algunas observaciones, 2) del lado derecho de la gráfica se muestra gráficamente la estrategia de separación del algoritmo SVM, el cual trata de centrarse dentro del espacio que hay entre clases, (elaboración propia).	23
2.6. Ejemplo de un modelo SVM para características bidimensionales donde se observan las características geométricas del algoritmo, (elaboración propia).	25
2.7. Unidad de umbral lineal (LTU, por sus siglas en inglés), (Crédito: Géron [30]).	26
2.8. Diagrama de un perceptrón, (Crédito: Géron [30]).	27
2.9. Diagrama de un perceptrón multicapa, (Crédito: Géron [30]).	28

2.10. Comparación de esquemas de selección de vecinos más cercanos; 1) Selección para el algoritmo Relief, 2) Selección para ReliefF, (Crédito: Urbanowicz [91])	32
4.1. Clasificación de Cobertura de Suelo Nacional derivada a partir de la aplicación de correspondencias entre las clases de CONABIO-INEGI sobre Serie V del Uso de Suelo y Vegetación publicada por el INEGI, resultando en las 12 clases usadas en el trabajo de Gebhardt y otros [28] y también en el presente estudio, (elaboración propia).	62
4.2. Identificación del porcentaje de píxeles de la clase mayoritaria para cada celda de 33 x 33 a partir de la Clasificación de Cobertura de Suelo Nacional resultante del mapeo de clases CONABIO-INEGI (elaboración propia).	65
4.3. Mapa de la selección aleatoria de 5,000 celdas por clase, en cada celda predomina al menos un 70 % de píxeles de una sola clase, dando como resultado un total de 60,000 regiones, (elaboración propia).	66
4.4. Distribución de 3,515 imágenes Landsat correspondientes al año 2011, se muestra el número de imágenes descargadas y procesadas para cada región, (elaboración propia).	67
4.5. Mosaico de la geomédiana Landsat nacional libre de nubes del año 2011, compuesta por píxeles de resolución de 30 metros x 30 metros, (elaboración propia).	70
4.6. Distribución de las clases de estudio a lo largo de Europa, en una muestra de 27,000 imágenes de 64 píxeles x 64 píxeles, (elaboración propia).	72
5.1. Proceso de transformación de un parche de imagen multi-banda a un renglón de características, (elaboración propia).	77
5.2. Representación visual de los 40 filtros de Gabor utilizados en el presente trabajo, (elaboración propia).	85
6.1. Histograma de distribución del mejor 10 % de características Landsat. Cada barra indica el porcentaje relativo al total de características de cada grupo por separado.	94

6.2. Histograma de distribución del mejor 10% de características Sentinel.	97
6.3. Histograma de la variación en la exactitud del clasificador Extra-Trees con muestras fijas para imágenes Landsat.	100
6.4. Histograma de la variación en la exactitud del clasificador MLP con muestras fijas para imágenes Sentinel.	101
6.5. Comparación de las estimaciones de calidad de las características; El eje X representa las estimaciones de calidad obtenidas con el método VLS-ReliefF y el Eje Y muestra algunas características enumeradas según la estimación de calidad lograda. 1) Ilustración de la estimación de calidad de las características Landsat, 2) Estimación de la calidad de las características Sentinel-2.	103
6.6. Comparación de la variación de los conjuntos de datos a través de los resultados obtenidos en 1,000 experimentos con el clasificador SVM. El eje X indica el número del experimento y el eje Y la exactitud lograda en los experimentos.	104

Índice de tablas

1.1. Índices espectrales identificados en este trabajo	2
1.2. Algunos proveedores de imágenes satelitales abiertas	3
3.1. Resumen de trabajos	52
4.1. Descripción de clases seleccionadas para este trabajo.	61
4.2. Correspondencias entre clases CONABIO-INEGI.	63
4.3. Porcentaje de mezcla de cada clase en el conjunto Landsat.	66
4.4. Lista de las bandas presentes en las imágenes Landsat	71
4.5. Áreas de entrenamiento y validación de Europa	72
5.1. Índices espectrales utilizados en la generación de características	78
5.2. Medidas estadísticas aplicadas a vectores numéricos en esta tesis	80
5.3. Medidas de textura estadísticas calculadas a partir del GLCM	81
5.4. Total de características por grupo y sensor	87
6.1. Desempeño de Landsat comparando el 10% de las mejores características vs Todas.	93
6.2. Distribución de las características y tiempos de procesamiento con imágenes Landsat	94
6.3. Exactitud por clase para el conjunto de datos Landsat, los resultados se alcanzaron con el algoritmo Extra-Trees.	95
6.4. Matriz de confusión de los resultados de la clasificación para Landsat, los resultados se alcanzaron con el algoritmo Extra-Trees.	95
6.5. Exactitud general de Landsat en comparación con el trabajo relacionado.	96
6.6. Desempeño de Sentinel-2 comparando el 10% de las mejores características vs Todas	96
6.7. Distribución de las características y tiempos de procesamiento con imágenes Sentinel-2	96

6.8. Exactitud por clase para el conjunto de datos Sentinel-2, resultados alcanzados con el algoritmo ANN.	98
6.9. Matriz de confusión de los resultados de la clasificación para Sentinel-2, resultados alcanzados con el algoritmo ANN.	98
6.10. Exactitud general de Sentinel-2 en comparación con el trabajo relacionado	98

Acrónimos y Abreviaciones

ARVI	Atmospherically Resistant Vegetation Index
BI	Bare Soil Index
CNN	Convolutional Neural Networks
CVUI	Coupled Vegetative Urban Index
DEM	Digital Elevation Model
EVI	Enhanced Vegetation Index
GCI	Green Chlorophyll Index
GCVI	Green Chlorophyll Vegetation Index
GEE	Google Earth Engine
GI	Green Index
GLCM	Gray-Level Co-Occurrence Matrix
GLDV	Gray-Level Difference Vector
GNDVI	Green Normalized Difference Vegetation Index
INEGI	Instituto Nacional de Estadística y Geografía
LDA	Linear Discriminant Analysis
LSWI	Land Surface Water Index
LULC	Land Use and Land Cover
MNDWI	Modified Normalized Difference Water Index
NDBI	Normalized Difference Built-up Index
NDMI	Normalized Difference Moisture Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near Infrared
OBIA	Object-based Image Analysis
ODC	Open Data Cube
OT	Observación de la Tierra

PCA	Principal Component Analysis
RF	Random Forest
RGRI	Red Green Ratio Index
SAVI	Soil Adjusted Vegetation Index
SRI	Simple Ratio Index
SRTM	Shuttle Radar Topography Mission
SVM	Support-Vector Machines
UI	Urban Index

Glosario

Análisis de componentes principales Técnica estadística que transforma un conjunto de variables correlacionadas en otras no correlacionadas con media cero. Que pueden escribirse como combinaciones lineales de las primeras.

Análisis discriminante lineal Método de clasificación supervisado en el que dos o más grupos son conocidos a priori y nuevas observaciones se clasifican en alguno de dichos grupos. Dada una nueva observación, se estima la probabilidad de pertenencia a alguno de los grupos haciendo uso del teorema de Bayes.

Áreas impermeables Áreas en las que el agua no puede penetrar al subsuelo.

Áreas urbanas Amplia región transformada por actividades humanas que generan impermeabilización del suelo típicamente debido al uso de materiales como asfalto y concreto. Lo anterior se realiza principalmente por la necesidad de nuevos espacios de vivienda.

Banda espectral Los sensores satelitales de percepción remota están diseñados para captar porciones del espectro electromagnético emitidos por la energía solar y reflejados por la superficie terrestre. A cada porción o rango dentro del espectro electromagnético se le conoce como banda espectral. Por ejemplo, los sensores Landsat 4, 5 y 7, tienen un sensor que capta la cantidad de energía en el rango de frecuencias de 0.45 a 0.52 micrómetros la cual corresponde a la región del color azul y de 0.52 a 0.60 micrómetros para la región del verde.

Bosque de coníferas También se le conoce como bosque templado porque presenta clima templado o semi-frío propio de las montañas. Según la especie que predomine, puede clasificarse como bosques de oyamel, ayarín, cedro, pino-encino, encino y abeto.

Bosque de manglar Bosques propios de las desembocaduras de cursos de agua dulce en latitudes tropicales y subtropicales. Así, entre las áreas con manglares se incluyen estuarios y zonas costeras; se caracterizan por árboles muy tolerantes a las sales existentes en dichas latitudes.

Bosque latifoliado Bosques propios de los climas cálidos y húmedos, templados y fríos; éstos se caracterizan por la presencia de árboles de hoja ancha como caoba, cedro, granadillo, nogal, etc.

Bosque mixto Son aquellos bosques que combinan bosques de hoja ancha (frondosos) de tipo caducifolio con bosques de coníferas, los cuales son de hoja acicular y perennifolios.

Campo de arroz Grandes extensiones de territorio donde se realiza la ricultura, es decir, el cultivo del arroz.

Campo de caña de azúcar Extensión territorial dedicada al cultivo de caña de azúcar.

Compuestos anuales y sub-anuales Imagen de satélite constituida a partir de una pila de imágenes correspondientes a periodos de tiempo variables, multi-anuales, anuales o sub-anuales. Cada píxel de la composición es resultado de un proceso de agregación, por ejemplo el píxel promedio de la pila de píxeles correspondiente al periodo de análisis.

Cuerpo de agua Acumulaciones de agua dulce permanentes que se encuentran sobre la superficie terrestre al interior de las áreas continentales.

DNB Corresponde a la banda pancromática día-noche del VIIRS la cual facilita la medición de luces nocturnas, luces solares / lunares reflejadas con un gran rango dinámico entre una iluminación mínima de un cuarto de luna y la luz del día más brillante.

Glaciar Masa de hielo y nieve originada en la superficie terrestre por acumulación y compactación.

Humedal Zona normalmente plana donde la superficie del suelo se cubre de agua de forma permanente o estacional, el sitio tiene que contener el líquido durante un período de tiempo suficiente para que prospere una población de plantas acuáticas.

Imagen de satélite Representación visual de los datos capturados por un sensor montado en un satélite artificial.

Índice espectral Combinaciones aritméticas de bandas espectrales que se construyen con la finalidad de aprovechar alguna propiedad física de los elementos de

la cobertura terrestre. Por ejemplo la vegetación refleja intensamente la luz infrarroja cercana.

Matorrales Comunidades vegetales dominadas por arbustos de altura inferior a 4 metros.

Misión Topográfica Radar Shuttle Proyecto internacional Agencia Nacional de Inteligencia-Geoespacial (NGA) y la Administración Nacional de Aeronáutica y el Espacio (NASA). La finalidad del proyecto fue adquirir datos radar desde el espacio para construir un mapa global de elevaciones topográficas.

MOD13Q1 Es un producto derivado de los satélites MODIS que se la NASA genera cada 16 días con el Índice Diferencial Normalizado de Vegetación (NDVI) y el Índice de Vegetación Mejorado (EVI) con la finalidad de que sean usados para generar comparaciones de vegetación a lo largo del tiempo.

Modelo digital de elevación Representación digital de los valores de altitud con respecto al nivel del mar que permite identificar el relieve de una región.

MODIS Son satélites desarrollados por la Administración Nacional de Aeronáutica y el Espacio (NASA) con resoluciones espaciales que van desde los 250 metros hasta los 1000 metros. Registra 36 bandas espectrales. Logra generar información de todo el planeta cada dos días.

Open Data Cube Solución de código abierto para el manejo de grandes volúmenes de imágenes de satélite listas para usarse.

Pastizales Son comunidades vegetales donde predominan los pastos con pocos árboles y arbustos. Pueden ser naturales o producto del desmonte de terrenos boscosos.

Percepción remota Se define como el arte, la ciencia y la tecnología a través de las cuales se identifican, miden y analizan las características de los objetos, ya sea sobre, encima o incluso debajo de la superficie terrestre, sin que exista un contacto directo entre los sensores y los objetivos o eventos que se observan.

Perceptrón multicapa Es una red neuronal que incluye una o varias capas ocultas totalmente conectadas.

Pradera Campo natural caracterizado por una cobertura de hierbas de me-

dia y corta altura, así como pequeños matorrales cuyo relieve es generalmente natural y ondulado con ríos entre sus colinas.

Random Forest Método de aprendizaje supervisado basado en la combinación de múltiples árboles de decisión que operan en conjunto para realizar las predicciones.

ResNet-50 Es una red neuronal de aprendizaje residual profundo para el reconocimiento de imágenes.

Sabana tropical Son praderas con árboles o arboledas dispersas que se encuentran en regiones cálidas con una o dos temporadas largas de sequía. La mayor extensión de tierra de este tipo se localiza en el centro y el este de África, aunque también hay grandes sabanas tropicales en Sudamérica y Australia.

Satélites LANDSAT Son una serie de satélites desarrollados por la Administración Nacional de Aeronáutica y el Espacio (NASA) con la finalidad de observar la superficie terrestre con sensores multiespectrales a una resolución media desde el año 1972. A lo largo de su historia han contado desde 7 hasta 11 bandas espectrales con una resolución espacial que va desde 15 metros en la banda pancromática, 30 metros en la mayoría de las bandas y 100 metros en bandas termales. Generan información del planeta cada 15 días.

Satélites Sentinel Son una serie de satélites de resolución media para la observación de la Tierra desarrollados por la Agencia Espacial Europea (ESA), entre los que se encuentran satélites radar (Sentinel-1) y satélites multiespectrales (Sentinel-2).

Support Vector Machines Método de aprendizaje supervisado que se basa en la construcción de hiperplanos para separar las clases del entrenamiento. El método busca un hiperplano que separe de forma óptima a los puntos de una clase de los de otra.

Tasseled cap transformation Método de transformación de la información espectral de las imágenes de satélite en indicadores espectrales. Dichos indicadores se enfocan en brillo, verdor y humedad.

Tierras agrícolas de riego Tierra dedicada a la agricultura que depende del uso de irrigación artificial.

Tierras agrícolas de temporal Tierra dedicada a la agricultura que depende de la temporada de lluvias para su irrigación.

Tierras desoladas Extensión de territorio sin cobertura vegetal. Puede tener roca desnuda o ser una zona desértica.

VIIRS Corresponde a un conjunto de sensores para la generación de imágenes infrarrojas visibles, instalados en uno de los satélites MODIS, las imágenes resultantes pueden ser aplicadas en detección de incendios y el monitoreo de la vegetación y océanos. Tiene 22 bandas espectrales con una resolución espacial de 371 metros y de 742 metros.



Capítulo 1
Introducción

Capítulo 1 Introducción

1.1. Motivación

El uso de datos de percepción remota se ha vuelto muy útil para generar información estadística sobre la sociedad y su entorno. En ese sentido, las técnicas de clasificación de uso y cobertura de suelo contribuyen a la identificación de las características de la superficie terrestre. En el proceso de toma de decisiones, la clasificación de la cobertura terrestre es relevante para complementarla con diversos tipos de información, como los generados por desastres naturales o procedimientos administrativos, como el mapeo catastral.

La literatura científica ha desarrollado numerosos índices espectrales, lo que representa una gran ventaja para el desarrollo de trabajos de análisis de percepción remota. La diversidad de índices permite una amplia gama de aplicaciones y permite la selección de aquellos índices más adecuados para las necesidades específicas de cada proyecto. En la tabla 1.1, se presentan algunos de los índices espectrales identificados durante el desarrollo del presente trabajo.

Como lo indica la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) [15] el término “cobertura de suelo”, se refiere al material físico en la superficie de la Tierra, por ejemplo bosque o desierto, diferenciándose así del término “uso de suelo”, el cual está definido por la actividad humana que se lleva a cabo en dicho territorio, por ejemplo agricultura o zona habitacional. Sin embargo, ambos están estrechamente relacionados, razón por la que muchos proyectos combinan la identificación de uso y la cobertura de suelo.

Llevar a cabo extracciones de información de uso y cobertura de suelo (LULC, por sus siglas en inglés de Land Use and Land Cover) se está convirtiendo en una tarea esencial para la ciencia, el monitoreo de recursos naturales y como una aproximación, para predecir datos sociodemográficos y económicos.

Tabla 1.1: Índices espectrales identificados en este trabajo

Índice	Referencia	Fórmula	Descripción
Atmospherically resistant vegetation index (ARVI)	[53]	$(NIR - (2 \times RED) + BLUE)/(NIR + (2 \times RED) + BLUE)$	Índice para el monitoreo de la vegetación con mejoras que disminuyen la interferencia de la atmósfera.
Built-up area extraction index (BAEI)	[9]	$(RED + 0.3)/(GREEN + SWIR1)$	Índice que combina las bandas Roja, Verde y SWIR1 para la detección de áreas urbanas.
Bare soil index (BI)	[96]	$[(SWIR1 + RED) - (NIR + BLUE)]/[(SWIR1 + RED) + (NIR + BLUE)]$	Índice para la detección de suelo descubierto.
Band ratio for built-up area (BRBA)	[97]	$RED/SWIR1$	Índice que combina las bandas Roja y SWIR1 para la detección de áreas urbanas.
Built-up Index (BU)	[104]	$[(SWIR1 - NIR)/(SWIR1 + NIR)] - [(NIR - RED)/(NIR + RED)]$	Índice que combina las bandas Roja, Infrarroja y SWIR1 para la detección de áreas urbanas.
Índice combinado de vegetación y urbanización (CVUI)	[94]	$[(SWIR2 - NIR)(NIR - RED)]/[(SWIR2 + NIR)(NIR + RED)]$	Índice para detectar la presencia de vegetación y áreas urbanas.
Enhanced Vegetation Index (EVI)	[40]	$2.5 \times (NIR - RED)/(NIR + 6 \times RED - 7.5 \times BLUE + 1)$	Índice de vegetación que considera múltiples condiciones atmosféricas y la cobertura de suelo bajo la vegetación.
Índice de Clorofila Verde (GCI)	[32]	$(NIR/GREEN) - 1$	Índice que se utiliza para estimar la cantidad de clorofila en la vegetación.
Índice de vegetación de clorofila verde (GCVI)	[63]	$(NIR/GREEN) - 1$	Índice para estimar la cantidad de clorofila en la vegetación.
Índice de verdor (GI)	[6]	$-0.2941(BLUE) + -0.243(GREEN) + -0.5424(RED) + 0.7276(NIR) + 0.0713(SWIR1) - 0.1608(SWIR2)$	Índice para detectar la salud y la densidad de la vegetación y la supervisión de sequías.
Índice de Vegetación de Diferencia Normalizada Verde-Infrarrojo cercano (GNDVI)	[33]	$(GREEN - NIR)/(GREEN + NIR)$	Índice que sirve para monitorear la vegetación, como en la agricultura y la silvicultura.
Index-based built-up index (IBI)	[102]	$(SWIR2 + 2 \times NIR + RED - GREEN)/(SWIR2 + 2 \times NIR + RED + GREEN)$	Índice que combina las bandas Roja, Verde, Infrarroja y SWIR2 para la detección de áreas urbanas.
Índice de Agua Modificado de Diferencia Normalizada (MNDWI)	[101]	$(GREEN - SWIR1)/(GREEN + SWIR1)$	Índice para identificar agua en áreas urbanas o rurales.
Normalized built-up area index (NBAI)	[97]	$(SWIR2 - SWIR1/GREEN)/(SWIR2 + SWIR1/GREEN)$	Índice que combina las bandas Verde, SWIR1 y SWIR2 para la detección de áreas urbanas y suelo descubierto.
New built-up index (NBI)	[51]	$RED \times SWIR1/NIR$	Índice combina las bandas Roja, Infrarroja y SWIR1 para la detección de áreas urbanas.
Índice de Diferencia de la Vegetación Normalizada de Construcciones (NDBI)	[104]	$(SWIR1 - NIR)/(SWIR1 + NIR)$	Índice para detectar la presencia de materiales construidos, como edificios y carreteras.
Índice de humedad de diferencia normalizada (NDMI)	[98]	$(NIR - SWIR1)/(NIR + SWIR1)$	Índice que refleja la cantidad de agua en las plantas y en la superficie del suelo.
Índice de vegetación de diferencia normalizada (NDVI)	[81]	$(NIR - RED)/(NIR + RED)$	Índice que mide la salud y densidad de la vegetación.
Índice de agua normalizado (NDWI)	[67]	$(GREEN - NIR)/(GREEN + NIR)$	Índice para detectar la humedad y la cantidad de agua en una superficie.
Índice de agua normalizado (NDWI2)	[80]	$(SWIR1 - RED)/(SWIR1 + RED)$	Índice para detectar la presencia de agua en la superficie terrestre.
Índice de la razón rojo verde (RGRI)	[26]	$RED/GREEN$	Índice que sirve para evaluar el estado de salud de la vegetación y para identificar áreas con problemas de crecimiento o estrés.
Índice de vegetación ajustado al suelo (SAVI)	[41]	$((NIR - RED)/(NIR + RED + 0.5)) \times (1.5)$	Índice para el monitoreo de la degradación de la tierra y en la evaluación de la calidad del suelo.
Simple ratio (SR)	[52]	NIR/RED	Índice que combina las bandas infrarroja y roja en una razón, útil para la detección de vegetación.
Urban Index (UI)	[54]	$(SWIR2 - NIR)/(SWIR2 + NIR)$	Índice que combina las bandas Infrarroja y SWIR2 para la detección de áreas urbanas.

Hoy en día, lo anterior es posible por el creciente número de sensores satelitales que proporcionan grandes volúmenes de datos de Observación de la Tierra (OT) [7,

20, 69, 89], y el aumento de datos satelitales de acceso gratuito como Landsat [99] y Sentinel-2 [90], que registran tiempos de 8 y 5 días de revisita y resoluciones de 30 y 10 metros con sus sensores multiespectrales [89]. Actualmente, los proveedores de servicios de cómputo y almacenamiento en la nube ofrecen acceso al acervo público de imágenes Landsat y Sentinel-2. En la Tabla 1.2 se muestran algunos de los servicios públicos de descarga de datos satelitales.

Tabla 1.2: Algunos proveedores de imágenes satelitales abiertas

Proveedor	Landsat	Sentinel-2
Amazon	https://registry.opendata.aws/landsat-8/	https://registry.opendata.aws/sentinel-2/
Google	https://cloud.google.com/storage/docs/public-datasets/landsat	https://cloud.google.com/storage/docs/public-datasets/sentinel-2

Por lo que ahora es posible combinar los datos abiertos que las instituciones públicas liberan con regularidad con las imágenes de satélite que se encuentran disponibles por múltiples medios. Esta gran disponibilidad de datos abiertos e imágenes satelitales abre posibilidades para usar estas fuentes de información y explotárlas a través del procesamiento de imágenes y técnicas de aprendizaje automático. Al mismo tiempo, la demanda de información para el monitoreo de los recursos naturales y las variables que pueden medirse en campo y buscarse una correlación estadística con las imágenes de satélite a través de técnicas de aprendizaje automático, se está volviendo más útil e interesante para muchas áreas del conocimiento.

Adicionalmente, es posible tener acceso a datos geográficos abiertos de alta calidad publicados por instituciones públicas como por ejemplo los que pone a disposición del público el Instituto Nacional de Estadística y Geografía (INEGI¹) en México.

Es importante realizar tareas como el análisis de la expansión de zonas urbanas, la generación de estadísticas agrícolas, información catastral, entre otras, para contribuir a los procesos de toma de decisiones basados en evidencia [73]. Se han realizado esfuerzos para identificar métodos de aprendizaje automático que contribuyan a la generación de estadísticas oficiales utilizando medidas adquiridas a través de datos de

¹<https://www.inegi.org.mx/>

observación de la tierra [4, 38, 60]. Por lo general, en muchos trabajos relacionados con la percepción remota, las características usadas en los procesos de aprendizaje automático son principalmente aquellas provenientes de valores espectrales o combinaciones de éstos en índices espectrales [46, 65, 79]. Trabajos más recientes han explorado el uso de la extracción de características con técnicas de procesamiento de imágenes digitales, extrayendo valores relacionados con el color, la textura o cualquier patrón que pueda extraerse de las imágenes para alimentar un método de aprendizaje automático y, por lo tanto, generar un modelo de clasificación [3, 62].

1.2. Descripción del Problema

El monitoreo de los recursos naturales depende de la generación de información estadística confiable y oportuna, sobre el estado y disponibilidad que dichos recursos naturales presentan.

En México, dicha generación de información se realiza por parte del INEGI mediante mapas que muestran la distribución del uso del suelo agrícola, de la vegetación natural e inducida del país, además indica el uso pecuario y forestal y otros usos que se presentan en el territorio relacionados con la cubierta vegetal [44]. A este tipo de información se le llaman colecciones LULC. La elaboración de estas colecciones de información se realiza mediante métodos manuales llevados a cabo por decenas de expertos en percepción remota, requiriendo un promedio de 5 años para llevar a cabo la metodología de actualización nacional de la clasificación de uso de suelo y vegetación.

La información de uso de suelo y vegetación que publica el INEGI, cuenta con 6 conjuntos de datos nacionales, según la Guía para la interpretación de cartografía del INEGI [1], se han realizado: Serie I (SI), elaborada entre 1979 y 1991; Serie II (SII), entre 1993 y 1999; Serie III (SIII), entre 2002 y 2005; Serie IV (SIV) entre 2006 y 2010; la Serie V (SV) entre 2011 y 2013 y Serie VI (SVI) entre 2014 y 2016. Las cuales usaron imágenes Landsat tomadas en los siguientes años:

- Serie I: Década de 1980
- Serie II: 1993

- Serie III: 2002
- Serie IV: 2007
- Serie V: 2011
- Serie VI: 2014

Se puede observar en los años de referencia de las imágenes que existen varios años entre las épocas de estudio, principalmente debido a que el proceso metodológico no ha recibido cambios radicales desde la serie II en 1993.

Debido al método usado, principalmente basado en trabajo de interpretación visual y verificaciones de campo, así como la cantidad de clases identificadas, prolonga el tiempo en el que se publica cada edición de la información de uso de suelo y vegetación. Teniendo un promedio de liberación de cinco años.

En este trabajo se buscan estrategias de clasificación automática que permitan acelerar los procesos de generación de conjuntos de información y contribuir para hacerlos más oportunos, aprovechando la información que ya se encuentra generada y clasificada. Es decir, aprovechar las colecciones de uso y cobertura de suelo actualmente disponibles para entrenar modelos que permitan acelerar los procesos de clasificación LULC y generar conjuntos de información mas rápidamente. Debido a que se requiere monitorear el cambio en superficie terrestre con más regularidad por la rapidez con la que se presentan los cambios derivados de la actividad humana y fenómenos globales que requieren datos más frecuentes en escalas nacionales. Aunque la metodología propuesta puede contribuir a la generación de colecciones de información sobre el uso y cobertura del suelo mas rápidamente, tiene como consecuencia la disminución del número de clases que se pueden identificar. Por lo que se considera que las técnicas de clasificación automática son complementarias al trabajo de fotointerpretación llevado a cabo por especialistas humanos. Los cuales pueden aprovechar los productos generados a mayor velocidad por parte de los métodos automatizados para detectar más rápidamente los cambios en la superficie terrestre y enfocarse a estudiar y generar información detallada en las zonas de mayor interés.

1.3. Objetivos

Objetivo General

Desarrollar una metodología para la clasificación automática de cobertura y uso de suelo mediante imágenes de satélite.

Objetivos Específicos

- Definir un método de extracción eficaz de características mediante técnicas de análisis de imágenes y reconocimiento de patrones.
- Explorar el impacto de diversos tipos de características en la exactitud global lograda en la clasificación LULC.
- Comparar múltiples clasificadores e identificar a los que tengan mejor desempeño en la tarea de clasificación LULC con imágenes de resolución media (Landsat y Sentinel-2).
- Comparar los resultados de las clasificaciones con trabajos del estado del arte.
- Generar una imagen multiespectral nacional libre de nubes del año 2011 para realizar el análisis nacional de cobertura de suelo.
- Publicar la base de datos de imágenes Landsat etiquetadas con las clases LULC utilizadas en este trabajo.

1.4. Contribuciones

En este manuscrito, se explora el uso de múltiples categorías de características derivadas a partir de los datos de las bandas espectrales y otras calculadas a partir de combinaciones bien conocidas en la literatura de percepción remota como índices espectrales, etc. Además de eso, se calcularon varias características a partir de cada imagen, por ejemplo, características basadas en texturas, filtros Gabor, varios resúmenes estadísticos de ellas, etc. Todas estas características fueron creadas para usarse como entrada para múltiples algoritmos que llevan a cabo la clasificación automática de

LULC. Además de esta exploración de características, se seleccionó la mejor combinación para alimentar un clasificador y obtener los mejores resultados posibles, y se realizó una comparación con trabajos relacionados.

Las pruebas de clasificación se realizaron con varios clasificadores bien conocidos, como Bosque Aleatorio (Random Forest), Máquinas de Vectores de Soporte (Support Vector Machines), Árboles Extremadamente Aleatorios (Extra-Trees) y Redes Neuronales Artificiales (Artificial Neural Networks), los cuales se usaron para comparar el rendimiento de diferentes tipos de características. Los experimentos de clasificación se realizaron partiendo las bases de datos en una proporción para entrenamiento y prueba de 70-30. Se realiza, además, un análisis de la relevancia de cada tipo de característica en relación a la clasificación, mostrando el desempeño de cada grupo de ellas en ambos tipos de imágenes utilizados. Con esta metodología, se han alcanzado resultados de desempeño comparables con trabajos relacionados del estado del arte (enfoques basados en el aprendizaje profundo y otros más tradicionales).

En resumen, las principales contribuciones de este documento son: el análisis comparativo de varios tipos de características para el problema de clasificación LULC, usando las bandas originales y bandas extendidas derivadas de índices calculados a partir de las originales, con lo que se extendieron las imágenes con las bandas derivadas de los índices espectrales calculados para cada píxel. A estas imágenes extendidas se les aplicaron métodos de caracterización mediante resúmenes estadísticos, así como rasgos de textura y forma. Se exploró el efecto de cada tipo de características comparando su desempeño con varios algoritmos de clasificación. En el caso de las imágenes Landsat se elaboró una base de datos que contiene 60,000 parches de imágenes Landsat etiquetados con 12 clases LULC, y está disponible para su uso y puede descargarse en la siguiente referencia [16]. El resultado de la clasificación de dichas imágenes es de 77.1% de exactitud la cual es una línea base que puede ser usada en otros trabajos de investigación relacionados con clasificación LULC. En el caso de Sentinel-2, se logró un 96.7% de exactitud global; los resultados obtenidos en este trabajo se compararon con cinco trabajos reportados recientemente en la literatura utilizando este mismo conjunto de datos. La metodología propuesta superó a cuatro de los cinco tra-

bajos relacionados, y el restante tiene un resultado muy cercano.

1.5. Publicaciones

A lo largo de la duración del doctorado se publicó el siguiente trabajo, la base de datos correspondiente y una imagen nacional libre de nubes.

Abel Coronado and Daniela Moctezuma “Feature evaluation for land use and land cover classification based on statistical, textural, and shape features over Landsat and Sentinel imagery”, *Journal of Applied Remote Sensing* 14(4), 048503 (16 November 2020). <https://doi.org/10.1117/1.JRS.14.048503>.

Abel Coronado and Daniela Moctezuma “Multispectral and augmented Landsat data with land cover labels (1.0) [Data set]”, Zenodo, (12 June 2020). <https://doi.org/10.5281/ZENODO.3891579>.

Abel Coronado and Daniela Moctezuma “Cloud free multispectral Landsat geomedian of Mexico for the year 2011 [Data set]”, Zenodo, (16 October 2022). <https://doi.org/10.5281/zenodo.7213479>.

1.6. Esquema de la Tesis

A continuación se describen los capítulos que componen la presente documento.

Capítulo : Introducción, se plantea la motivación que llevó a la realización del presente trabajo así como el planteamiento del problema, objetivos y contribuciones.

Capítulo 1.6: Fundamentos, en esta sección se realizó una revisión de los conceptos básicos sobre percepción remota y aprendizaje automático, requeridos para la elaboración del presente trabajo.

Capítulo 2.4: Trabajos Relacionados, en este capítulo se describen los trabajos

en que se realizan investigaciones similares a esta tesis, identificando las distintas estrategias de extracción de características y los resultados al aplicar diversos métodos de clasificación a imágenes de satélite.

Capítulo 3.2: Bases de Datos, en este capítulo se describen las características de la base de datos usada en el caso de las imágenes Sentinel-2 y el procedimiento llevado a cabo para construir la base de datos de imágenes Landsat.

Capítulo 4.3: Clasificación de cobertura y uso de suelo, en este capítulo se describe la metodología para la extracción de características a partir de las imágenes, así como el proceso de clasificación automática de cobertura y uso de suelo utilizando imágenes de satélite.

Capítulo 5.7: Experimentos y Resultados, en este capítulo se analizan los resultados de los experimentos realizados en este manuscrito, identificando las combinaciones de características que más contribuyen para obtener resultados destacados en comparación con otras estrategias de extracción de características y clasificación.

Conclusiones y trabajo futuro: En este capítulo se realiza una síntesis breve de los aspectos más relevantes y aportaciones de la presente investigación. También exploramos algunas ideas del potencial trabajo futuro derivado del trabajo actual.



Capítulo 2

Fundamentos

Capítulo 2 Fundamentos

En este documento llevaremos acabo el aprovechamiento de múltiples conceptos acerca de la percepción remota, procesamiento digital de imágenes y aprendizaje automático (machine learning). Por lo que en esta sección realizaremos una revisión de los conceptos básicos requeridos para llevar a cabo los trabajos reportados en los siguientes capítulos.

2.1. Percepción Remota

En el desarrollo de los trabajos de esta tesis entendemos percepción remota como la práctica de obtener información de la superficie terrestre usando imágenes satélites, generadas a partir de la radiación de una o mas regiones del espectro electromagnético reflejada o emitida desde la superficie terrestre [13].

De acuerdo al programa de entrenamiento de percepción remota aplicada (AR-SET, por sus siglas en inglés) de la NASA, tenemos los siguientes conceptos básicos:

La percepción remota satelital se lleva a cabo mediante instrumentos o sensores que miden la radiación electromagnética emitida por el sol y que es reflejada por la tierra, estos sistemas se denominan "sistemas de percepción remota pasivos". La intensidad de la radiación reflejada y emitida al espacio es influenciada por las condiciones de la superficie y la atmósfera. Por lo tanto, las mediciones satelitales contienen información sobre estas condiciones [5].

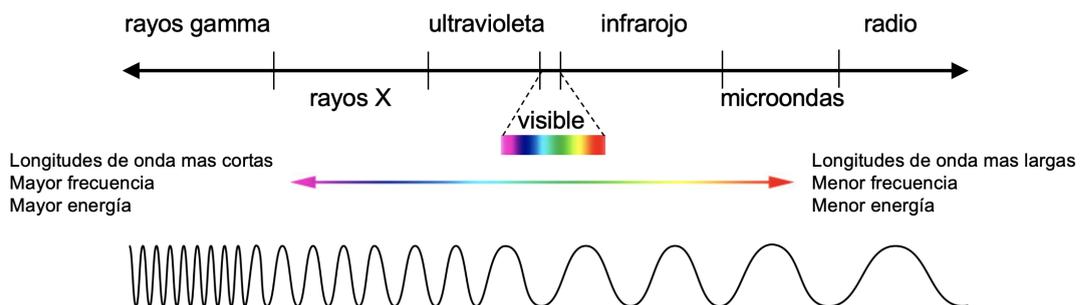


Figura 2.1: Comparación de la longitud de onda, frecuencia y energía del espectro electromagnético (crédito: imagina el universo de la NASA).

En la figura 2.1 se muestran las longitudes de onda emitidas por el sol las cuales van desde longitudes de varios kilómetros, conocidas como ondas de radio, hasta longitudes pequeñísimas en el rango de los micrómetros, conocidas como rayos gamma. La atmósfera filtra los rayos gamma, rayos X y una parte de la radiación ultravioleta; y permite el paso de una parte de la radiación ultravioleta, luz visible y radiación infrarroja. Los sensores de los satélites están diseñados para cuantificar la intensidad de energía que les llega en ciertas longitudes de onda. Por ejemplo, el satélite Landsat 8 tiene un sensor que capta las longitudes de onda entre 0.53 y 0.59 micrómetros correspondientes a las emitidas por el color verde de la luz visible. A estas regiones de observación se les llama bandas espectrales.

La energía irradiada por el sol interactúa con los objetos en la superficie terrestre; por ejemplo, cuando la radiación llega a la vegetación sana, ésta absorbe las longitudes de onda correspondientes a los colores azul y rojo reflejando un poco del color verde y prácticamente todas las ondas infrarrojas. En el caso de cuerpos de agua las longitudes de onda correspondientes al rojo e infrarrojo son absorbidas reflejando más cantidad de ondas como el azul y algo de verde.

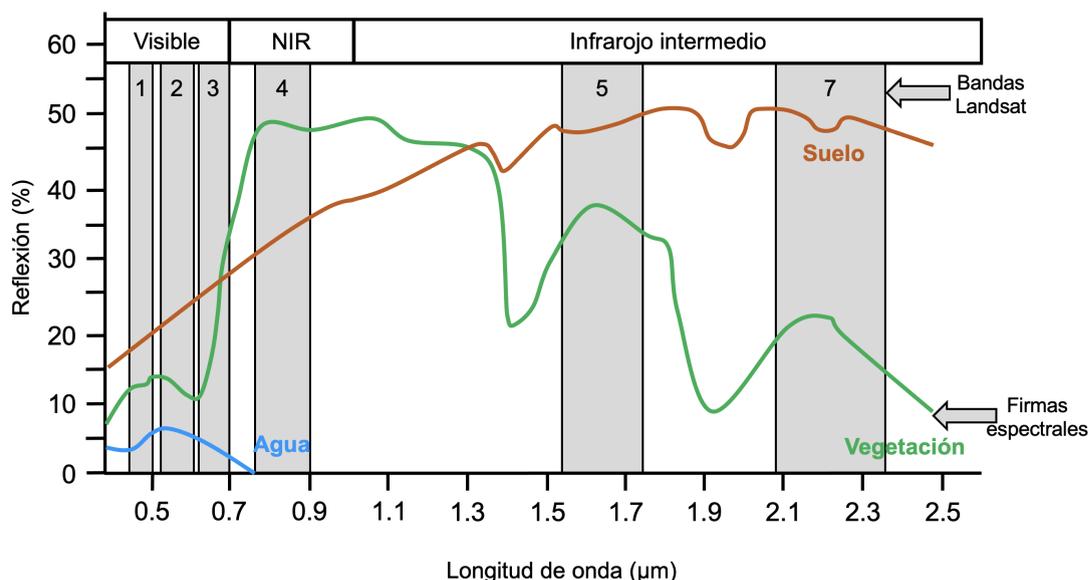


Figura 2.2: Ejemplos de firmas espectrales. Las franjas corresponden a las bandas espectrales Landsat: 1 azul, 2 verde, 3 rojo, 4 infrarrojo, 5 SWIR 1 y 7 SWIR 2 (Crédito: ARSET de la NASA).

Cada tipo de superficie tiene una manera única de reflejar el espectro electro-

magnético, como se aprecia en la figura 2.2 obtenida de ARSET de la NASA [5]. Se pueden observar tres ejemplos de firmas espectrales; en la línea color café, podemos ver la respuesta promedio del suelo desnudo que refleja primordialmente el infrarrojo intermedio captado por las bandas SWIR1 y SWIR 2; en la línea color verde oscuro, se muestra la firma de la vegetación sana con un leve incremento en la banda verde y una marcada intensidad en la banda infrarroja y por último, la línea azul que representa la respuesta espectral del agua que refleja principalmente verde y azul con una respuesta baja en la banda roja y absorbe prácticamente todas las longitudes de onda infrarrojas.

2.1.1. Características de los satélites y sus sensores

La información de los satélites utilizados en éste trabajo se obtiene gracias a recorridos mediante órbitas circulares de movimiento constante a una distancia relativa de la tierra que va desde 160-2000 km. La órbita de los satélites Landsat y Sentinel-2 es polar, es decir, es una órbita que pasa por encima de los polos.

Dichas órbitas se caracterizan por una cobertura global con una frecuencia de las mediciones que van desde 1 a 30 días. En el caso de los satélites utilizados en este manuscrito, corresponden a la órbita polar y la frecuencia temporal, también conocida como resolución temporal, es de 8 días para los satélites Landsat combinados y 5 días para los satélites Sentinel-2.

El tipo de sensor utilizado por los satélites Sentinel-2 y Landsat es denominado sensor pasivo, los cuales miden la energía reflejada o emitida por la tierra y su atmósfera. La energía radiante se convierte a cantidades biofísicas tales como temperatura, precipitación, humedad del suelo, clorofila, etc.

Se define como resolución espectral a la capacidad del sensor para captar diferentes longitudes de onda. Los sensores correspondientes a los satélites utilizados en este trabajo son sensores multiespectrales. En el caso de los sensores Landsat captan información de siete a 11 bandas espectrales y en el caso de Sentinel-2 capta información de 13 bandas espectrales. Existen otros instrumentos denominados hiperespectrales como el sensor AVIRIS que logra captar 256 bandas espectrales.

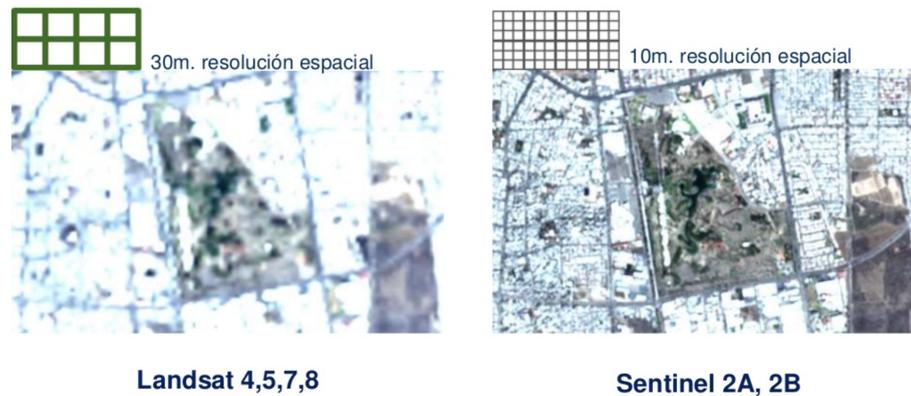


Figura 2.3: Comparación de la resolución espacial de los satélites Landsat y Sentinel-2, (elaboración propia).

La resolución espacial se determina a partir de la unidad más pequeña que un sensor puede medir, correspondiente a la cobertura espacial en la superficie terrestre. Es decir, se refiere al máximo detalle que se puede distinguir por los píxeles que conforman a una imagen. En la imagen 2.3, de elaboración propia, se puede observar la comparación entre la resolución espacial del satélite Landsat a la izquierda y del satélite Sentinel-2 a la derecha. La retícula que se encuentra en la parte superior izquierda de cada imagen representa una comparación entre los tamaños de cada píxel en la superficie terrestre. Los píxeles de Sentinel-2 pueden captar nueve veces más detalle de la superficie que los píxeles Landsat.

Por otro lado, la resolución radiométrica se refiere al número máximo de niveles de luminosidad disponibles, lo cual depende de la cantidad de bits utilizados para la representación de la energía registrada por el sensor. Caso de Landsat la resolución radiométrica es de 16 bits en el de Sentinel-2 es de 12 bits.

2.1.2. Características de las observaciones de percepción remota

Las observaciones de percepción remota ofrecen información global y consistente aún cuando no se cuenten con mediciones a nivel de suelo. Aunque no proporcionan un alto nivel de detalle a nivel de suelo, por ejemplo, no pueden detectar la composición del manto terrestre bajo el dosel forestal y se encuentran limitadas en el nivel de profundidad capaz de detectar bajo el agua.

Es muy difícil obtener observaciones con altas resoluciones espaciales espectrales, temporales y radiométricas al mismo tiempo. Existen satélites con la capacidad de generar coberturas globales diariamente; sin embargo, gracias a su amplia cobertura de barrido carecen de una buena resolución espacial. Los satélites de órbita polar pueden generar imágenes de más alta resolución pero pueden tardar de cinco a 16 días en realizar una cobertura global. Por otro lado, los satélites geostacionarios obtienen imágenes con mucha mayor frecuencia sin embargo, la resolución espacial es menor debido a que la distancia orbital es mucho mayor: se encuentran más alejados de la tierra.

Las observaciones de percepción remota se caracterizan por tener múltiples formatos y características, por lo que puede llegar a ser complejo el manejo de la información satelital. Las aplicaciones de estos tipos de datos pueden requerir procesamientos, herramientas y visualizaciones adicionales.

2.1.3. Niveles de procesamiento de los datos satelitales

Los niveles de procesamiento corresponden a la cantidad de procesamiento posterior que se lleva a cabo en las imágenes satelitales. El nivel cero de procesamiento corresponde a las imágenes salidas directamente del sensor satelital sin ningún tipo de corrección adicional. Corresponden a los datos brutos obtenidos directamente del satélite. Este tipo de datos es el más difícil de utilizar debido a que se deben considerar múltiples etapas de corrección y procesamiento para poder hacerlos útiles y comparables. El primer nivel de procesamiento implica correcciones geométricas, a partir de puntos de control terrestres, dando como resultado datos calibrados y geocalizados. El segundo nivel de procesamiento corresponde a productos derivados del primero, como lo puede ser el cálculo de variables e indicadores geofísicos. El tercer nivel de procesamiento corresponde a los datos llamados Analysis Ready Data (ARD), se procesan constantemente con los más altos estándares científicos y el nivel de procesamiento requerido para su uso directo en el monitoreo y la evaluación de cambios en el paisaje. Los ARD reducen significativamente la magnitud del pre-procesamiento requerido por los datos, eliminando la necesidad de descargar y preparar grandes can-

tidades de datos para análisis de investigación de series de tiempo. Los datos Landsat usados en este documento corresponden el tercer nivel de procesamiento.

2.2. Fundamentos de aprendizaje automático

De acuerdo con los autores de los libros [8, 11, 47] el aprendizaje automático o Machine Learning es un subcampo de las ciencias de la computación que se encarga de diseñar algoritmos que aprenden a partir de un conjunto de datos que caracterizan algún fenómeno. Éstos ejemplos pueden provenir de la naturaleza, ser construidos por humanos o generados por otro algoritmo.

Existen varios tipos de aprendizaje automático, en esta investigación nos centraremos en el aprendizaje supervisado. Como lo indica Burkov [11] en el aprendizaje supervisado el conjunto de datos corresponde a una colección de ejemplos o instancias que tienen asignada una etiqueta que los categoriza y del cual se extraen aleatoriamente conjuntos de entrenamiento y prueba. Cada elemento del conjunto de datos mencionado se denomina vector de características. Un vector de características es un vector en el que cada dimensión contiene un valor que describe al objeto de estudio de alguna manera. Por ejemplo, en el caso de las imágenes de satélite, cada ejemplo puede representar a una imagen etiquetada, la primera característica podría tener el promedio de todos los valores de la banda espectral correspondiente a la longitud de onda del rojo, la segunda puede ser el promedio de la banda azul y la tercera podría ser el promedio de la banda verde y así sucesivamente para todos los ejemplos en el conjunto de datos. La característica ubicada en cada columna corresponde siempre al mismo tipo de información. Esto significa que si la segunda columna contiene el promedio de la banda azul en los ejemplos posteriores del conjunto de datos la segunda columna también corresponderá al promedio de la banda azul. La etiqueta puede ser un elemento perteneciente a un conjunto finito de clases o un número real, también puede ser una estructura más compleja como un vector, una matriz, un árbol o un grafo. Una clase puede verse como una categoría a la que pertenece un ejemplo, es decir, si los ejemplos son imágenes de satélite y el problema es la detección de clases de cobertura de suelo y uso de suelo, las clases podrían ser bosque, desierto o ciudad.

El objetivo de un algoritmo de aprendizaje supervisado es usar el conjunto de datos para producir un modelo que tome un vector de características como datos de entrada y genera como salida la información que permite deducir la etiqueta de dicho vector de características. Por ejemplo, el modelo creado con un conjunto de datos de imágenes de satélite podría tomar como entrada un vector de características que describa a las imágenes y como salida generar una probabilidad de que la imagen corresponda a la clase bosque o ciudad.

2.2.1. Algoritmo de clasificación supervisada

En esta sección haremos una revisión de las características principales de los algoritmos de clasificación supervisada utilizados para la realización de los experimentos de este trabajo. Iniciaremos nuestra revisión con los métodos basados en árboles, de los cuales analizaremos a los árboles de decisión, el algoritmo de Bosque Aleatorio (Random Forest) y Árboles Extremadamente Aleatorios (Extra-Trees). Posteriormente analizaremos los fundamentos de las Máquinas de Vectores de Soporte (Support Vector Machines) y concluiremos con el Perceptrón Multicapa.

Métodos basados en arboles

Los métodos basados en árboles son utilizados en problemas de clasificación supervisada y regresión. Éstos métodos consisten en la segmentación recursiva del espacio de características en regiones cuboidales cuyos vértices están alineados a los ejes correspondientes de las dimensiones de dicho espacio de características. Un árbol de clasificación es el resultado de la secuencia ordenada de evaluaciones simples sobre una sola característica a la vez. La secuencia de preguntas concluye en la predicción de una clase.

Considerando un ejemplo simple, del lado izquierdo de la figura 2.4 se tiene la segmentación recursiva binaria de un espacio de características para dos variables de entrada, x_1 y x_2 . Del lado derecho de la imagen se tiene la estructura del árbol binario correspondiente. Las posibles etapas serían las siguientes: (1) ¿Es $x_1 \leq \theta_1$? si la respuesta es sí, se sigue la rama izquierda; si no, se sigue la rama derecha. Si la respuesta a (1) es

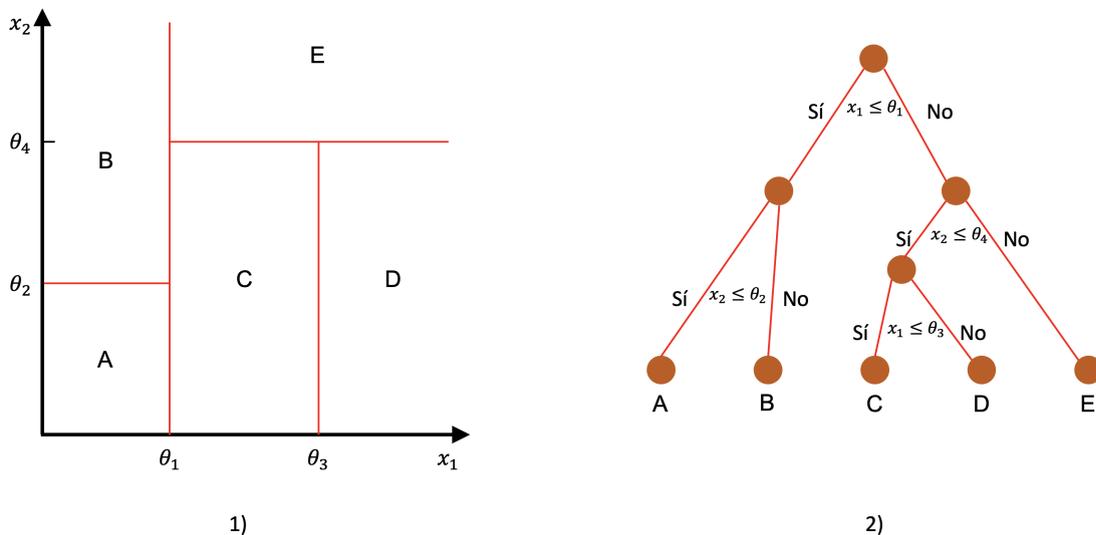


Figura 2.4: Ejemplo de segmentación del espacio de características usando decisiones binarias; 1) Segmentación del espacio de características a partir de múltiples umbrales, 2) Árbol de decisión correspondiente, el cual muestra la secuencia de preguntas binarias sobre el espacio de características, (elaboración propia).

si entonces la siguiente pregunta es (2): ¿ $x_2 \leq \theta_2$? si la respuesta es sí entonces se termina en el nodo hoja **A** correspondiente a la región $A = \{x_1 \leq \theta_1, x_2 \leq \theta_2\}$; si la respuesta es no entonces va al nodo terminal **B** correspondiente a la región $B = \{x_1 \leq \theta_1, x_2 > \theta_2\}$. (3) si la respuesta a (1) es no, hacemos la siguiente pregunta ¿Es $x_2 \leq \theta_4$? Si la respuesta a (3) es sí, entonces se sigue por la rama izquierda y se hace la siguiente pregunta (4): ¿Es $x_1 \leq \theta_3$? si la respuesta es si entonces se produce el nodo terminal **C** correspondiente a la región $C = \{\theta_1 < x_1 \leq \theta_3, x_2 \leq \theta_4\}$; si la respuesta a (4) es no llegamos al nodo terminal **D** en la región $D = \{\theta_1 < x_1 > \theta_3, x_2 \leq \theta_4\}$. Por ultimo si la respuesta a (3) es no, llegamos al nodo terminal **E** correspondiente a la región $E = \{x_1 > \theta_1, x_2 > \theta_4\}$. La partición en 5 regiones resultante se muestra en el lado izquierdo de la figura 2.4. Para un árbol de clasificación cada nodo terminal y región correspondiente tiene asignada una clase.

Un atributo clave de los modelos de clasificación basados en árboles de decisión es que son fácilmente interpretables por humanos porque corresponden a una secuencia de decisiones binarias aplicadas a variables de entrada individuales. Para poder identificar este tipo de modelos a partir de un conjunto de entrenamiento, es necesario determinar la estructura del árbol, incluyendo cuál variable de entrada será elegida en cada nodo para formar el criterio de segmentación, así como también

determinar el umbral del parámetro θ_i para una nueva división del espacio de características.

De acuerdo con Bishop [8], el problema de determinar la estructura óptima de un árbol de decisión mediante una técnica analítica de optimización, suele ser computacionalmente inviable debido al gran número de combinaciones de las posibles soluciones. En lugar de eso, se suele implementar una estrategia de optimización voraz empezando con un solo nodo raíz correspondiente a todo el espacio de entrada, entonces se crece el árbol agregando un nodo a la vez. A cada paso habrá algún número de regiones candidatas en el espacio de entrada que podrán ser divididas, correspondiente a la adición de un par de nodos hoja al árbol existente. Para cada uno de esos nodos hoja existe la elección de cuál de las variables de entrada se utilizará para la división así como también el valor del umbral. La optimización conjunta de la elección para la división de la región consiste en elegir la variable de entrada y el umbral, lo cual puede realizarse eficientemente mediante una búsqueda exhaustiva considerando que, para una variable de decisión y un umbral, dados la selección óptima de la variable predictiva, está dada por el promedio local de los datos para cada etiqueta. Esto se repite para todas las posibles elecciones de variable para la segmentación y aquella con la suma del cuadrado del error residual menor es retenida. La estrategia de optimización voraz para crecer el árbol mantiene el problema de cuándo detener el incremento de nodos. Una estrategia simple puede ser detenerse cuando la reducción en el error residual se encuentra dentro de un umbral. Sin embargo, se ha encontrado empíricamente que ninguna de las divisiones disponibles produce una reducción significativa del error. Por esta razón, es una práctica común crecer un árbol grande usando un criterio de paro basado en el número de puntos asociados con los nodos hoja, y entonces podar el árbol resultante.

La poda del árbol se basa en un criterio que balancea los errores residuales contra una medida de complejidad del modelo. En el caso de los árboles de clasificación se utiliza una medida de desempeño, donde se toma en cuenta la proporción de observaciones de entrenamiento en cada región resultante asignados a cada clase. Dos elecciones comunes son el índice de Gini y la entropía cruzada. El índice de Gini se

define como:

$$G_i = 1 - \sum_{k=1}^K p_{i,k}^2 \quad ,$$

Dónde $p_{i,k}$ es la proporción de la cantidad de ejemplos de la clase k del conjunto de entrenamiento en el i -ésimo nodo. Es fácil ver que el índice de Gini toma valores cercanos a uno si las $p_{i,k}$'s son pequeñas, y toma valores cercanos a cero si una $p_{i,k}$ es cercana a uno. Por esta razón el índice de Gini se refiere como una medida de la pureza del nodo. Por otra parte, la entropía cruzada se define como:

$$E_i = - \sum_{k=1}^K p_{i,k} \log p_{i,k} \quad .$$

Se puede mostrar que la entropía cruzada tomará un valor cercano a cero si todas las $p_{i,k}$'s son muy similares entre sí. Por lo que el índice de Gini y la entropía cruzada son muy similares numéricamente; por lo que pueden ser usadas como medidas de pureza del nodo.

La interpretabilidad de los modelos de clasificación basados en árboles es vista como su mayor fortaleza. Sin embargo, en la práctica se ha encontrado que una estructura particular derivada de un conjunto de datos de entrenamiento es muy sensible a dicho conjunto de datos de tal manera que un pequeño cambio en dicho conjunto resulta en segmentaciones totalmente diferentes.

Hay otros problemas con los métodos de clasificación basados en árboles, uno de los cuál es que los segmentos están alineados a los ejes del espacio de características, lo cual puede ser muy ineficiente. Por ejemplo para separar dos clases cuya frontera de decisión está en los 45° requerirá un gran número de divisiones paralelas a los ejes del espacio de entrada comparado a una simple división que no esté alineada a los ejes.

Los métodos basados en árboles de decisión son simples y fácilmente interpretables. Sin embargo, no son tan competitivos con respecto a otros algoritmos de aprendizaje supervisado.

Para atenuar el impacto de la gran variabilidad que presentan los árboles de decisión comentados arriba, se aplica un procedimiento de propósito general para reducir la varianza. Dicho procedimiento es denominado *bagging* (del inglés: Bootstrap Aggregating). El método consiste en tomar múltiples muestras a partir del conjunto de entrenamiento, con esta estrategia podemos generar múltiples conjuntos de entrenamiento los cuáles son resultado del re-muestreo mencionado. Aplicar *bagging* a métodos basados en árboles es muy útil para mejorar sus predicciones. Desafortunadamente la mejora lograda en la exactitud de la predicción afecta la interpretabilidad del modelo resultante, debido a que ahora ya no es posible representar el modelo en un solo árbol de decisión.

Bosque Aleatorio (Random Forest)

El algoritmo de clasificación llamado Random Forest provee una mejora sobre los árboles de decisión con la técnica de re-muestreo llamada *bagging*. Al igual que en dicho método, se construyen múltiples árboles de decisión a partir de múltiples muestras de entrenamiento (*bootstrapping*). Adicionalmente, cuando se construyen esos árboles de decisión, se lleva a cabo una división del espacio de características considerando una muestra aleatoria de características como variables candidatas extraídas a partir del conjunto completo de variables predictoras. Lo cual se repite nuevamente para cada segmentación. En otras palabras, al construir un modelo de Random Forest cada división del espacio de características no se consideran en la mayoría de las variables predictoras o características. Lo cual da como resultado árboles con menor variabilidad y más confiables debido a estas estrategias de segmentación aleatoria.

Árboles Extremadamente Aleatorios (Extra-Trees)

Cuándo se construye un modelo de clasificación basado en árboles usando el algoritmo Random Forest, cada nodo solamente considera un subconjunto aleatorio de características para realizar la segmentación. Es posible construir árboles aún más aleatorios usando umbrales aleatorios para cada característica en lugar de buscar el mejor umbral posible. Un bosque de tales árboles aleatorios es llamado ensamble de árboles extremadamente aleatorios, también conocidos como Extra-Trees. Al igual que

en el algoritmo de Random Forest, se obtiene mayor sesgo a cambio de menor varianza. Esto hace a los Extra-Trees más rápidos de entrenar que los Random Forest dado que encontrar el mejor umbral posible para cada variable predictora es una de las tareas que más tiempo consumen al crecer un árbol.

Máquinas de Vectores de Soporte (Support Vector Machines)

Según James et al. [48] las máquinas de vectores de soporte (SVM, por sus siglas en inglés de Support Vector Machine), se consideran frecuentemente uno de los mejores algoritmos de clasificación. También es destacable su alto desempeño en múltiples situaciones, incluso cuando se lidia con grandes volúmenes de información. La idea fundamental detrás de SVM se puede explicar gráficamente en la figura 2.5. Los datos representados en las gráficas muestran dos clases (una en color azul y la otra en color naranja) claramente separables mediante una línea recta (linealmente separables). El gráfico de la izquierda muestra las fronteras de decisión de tres posibles clasificadores lineales en los que se puede ver que los modelos se encuentran muy cerca de algunas observaciones lo cual probablemente no se desempeñe adecuadamente con nuevos ejemplos. En contraste, la línea sólida de la gráfica del lado derecho representa la frontera de decisión de un clasificador SVM, la línea no solamente separa claramente las dos clases sino que también se encuentra tan lejos como es posible de las instancias de entrenamiento. Se puede pensar al clasificador SVM como el ajuste de la carretera más ancha posible representada por las líneas paralelas punteadas entre las dos clases. Se puede observar que, agregar más instancias de entrenamiento fuera de la región marcada por la carretera no afecta la frontera de la decisión en absoluto, debido a que está totalmente determinada o soportada por los ejemplos ubicados en la orilla de la carretera. A esas instancias se les llama vectores de soporte y se encuentran identificados mediante un círculo.

Como explica Burkov [11], para definir la frontera de decisión del algoritmo SVM se utiliza la ecuación del hiperplano que está dada por dos parámetros, un vector de valores reales w de la misma dimensión que el vector de características de entrada

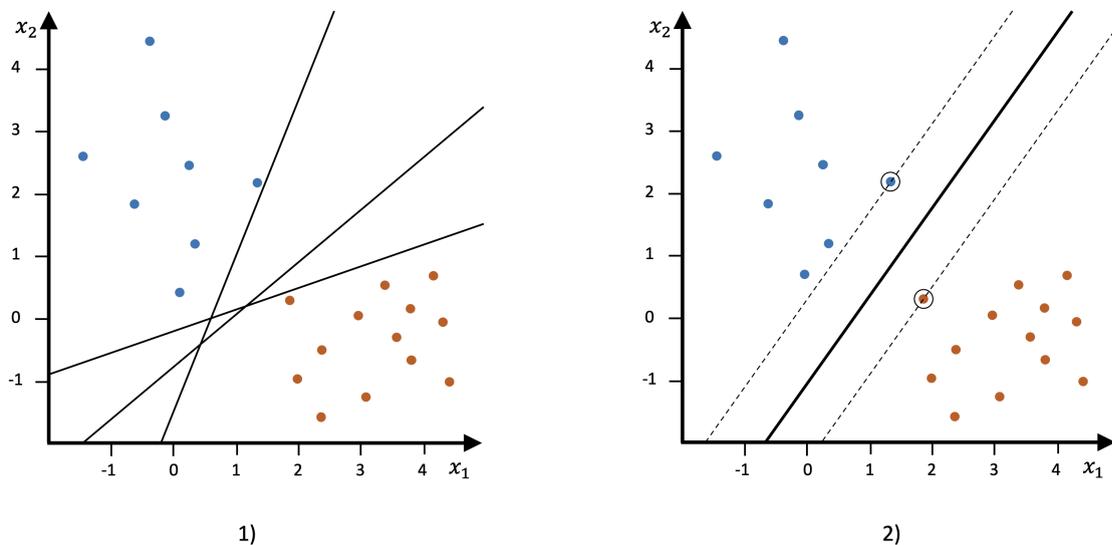


Figura 2.5: Ejemplo de segmentación del espacio de características límites de decisión: 1) del lado izquierdo se muestran múltiples fronteras de decisión donde se puede ver que están muy pegadas a algunas observaciones, 2) del lado derecho de la gráfica se muestra gráficamente la estrategia de separación del algoritmo SVM, el cual trata de centrarse dentro del espacio que hay entre clases, (elaboración propia).

\mathbf{x} , y un número real b de tal manera que la frontera de decisión se define por:

$$\mathbf{w}\mathbf{x} - b = 0,$$

Donde $\mathbf{w}\mathbf{x}$ significa $w_1x_1 + w_2x_2 + \dots + w_Dx_D$, y D es el número de dimensiones del vector de características \mathbf{x} .

En el caso de la clasificación binaria, la etiqueta a predecir a partir de un vector de entrada \mathbf{x} viene dada de la siguiente forma:

$$y = \text{sign}(\mathbf{w}\mathbf{x} - b),$$

En el contexto de los algoritmos SVM sign es una función matemática que toma cualquier valor como entrada y devuelve +1 si la entrada es un número positivo o -1 si la entrada es un número negativo.

El objetivo del algoritmo de aprendizaje SVM es usar el conjunto de datos de entrenamiento y encontrar para los parámetros \mathbf{w} y b los valores óptimos \mathbf{w}^* y b^* .

Una vez que el algoritmo de aprendizaje identifica estos valores óptimos, el modelo $f(\mathbf{x})$ se define como:

$$f(\mathbf{x}) = \text{sign}(\mathbf{w}^* \mathbf{x} - b^*).$$

Ahora para que el algoritmo encuentre w^* y b^* requiere resolver un problema de optimización bajo ciertas restricciones:

$$\mathbf{w}x_i - b \geq +1 \quad \text{si } y_i = +1,$$

$$\mathbf{w}x_i - b \leq -1 \quad \text{si } y_i = -1,$$

También es preferible que el hiperplano separe los ejemplos positivos de los negativos con el mayor margen posible. El margen es la distancia entre los ejemplos más cercanos de las dos clases, tal y como vienen definidos por la frontera de decisión. Un amplio margen contribuye a una mejor generalización, que es como el modelo clasificará a nuevos ejemplos en el futuro. Si buscamos el hiperplano separador óptimo con los márgenes más amplios posibles entre las dos clases, necesitamos minimizar la norma euclidiana de \mathbf{w} denotada por $\|\mathbf{w}\|$ y dada por $\sqrt{\sum_{j=1}^D w_j^2}$

Para vectores de características bidimensionales, el problema de la solución se pueden visualizar en la figura 2.6. Los círculos azul y naranja representan, respectivamente, ejemplos positivos y negativos de la línea dada por $\mathbf{w}\mathbf{x} - b = 0$ que es la frontera de decisión.

Geoméricamente, las ecuaciones $\mathbf{w}\mathbf{x} - b = 1$ y $\mathbf{w}\mathbf{x} - b = -1$ definen dos hiperplanos paralelos, como se ve en la figura 2.6. Se puede mostrar que la distancia entre esos planos está dada por $\frac{2}{\|\mathbf{w}\|}$ así que, cuanto menor sea la norma $\|\mathbf{w}\|$ mayor será la distancia entre los hiperplanos.

Así es como funcionan las Support Vector Machines. Esta versión particular del algoritmo construye lo que se conoce como un modelo lineal. Se llama lineal porque la frontera de decisión es una línea recta (un plano, o un hiperplano). Las SVM también pueden incorporar kernels que pueden hacer que la frontera de decisión sea arbitrariamente no lineal. En algunos casos será imposible separar perfectamente los dos grupos

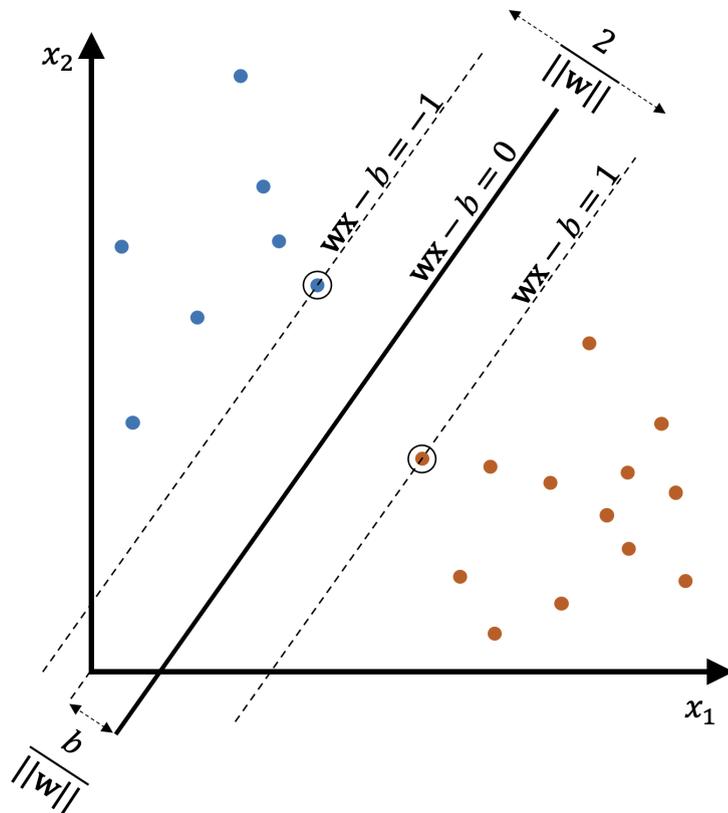


Figura 2.6: Ejemplo de un modelo SVM para características bidimensionales donde se observan las características geométricas del algoritmo, (elaboración propia).

debido al ruido en los datos, errores en el etiquetado, o valores atípicos. Otra versión de SVM puede incorporar un hiperparámetro de penalización para la clasificación errónea de ejemplos de entrenamiento de clases específicas.

Redes Neuronales (Artificial Neural Networks)

De acuerdo al texto del libro de Géron [30] el perceptrón es una de las arquitecturas de redes neuronales artificiales (ANN, de sus siglas en inglés de Artificial Neural Networks) más simples. Está basada en una neurona artificial llamada unidad de umbral lineal (LTU, por sus siglas en inglés Linear Threshold Unit). Las entradas y salidas son numéricas, cada entrada está asociada con un peso. La LTU calcula la suma ponderada de sus entradas ($z = w_1x_1 + w_2x_2 + \dots + w_Dx_D = \mathbf{w}\mathbf{x}$) entonces se aplica la función step: $h_w(\mathbf{x}) = \text{step}(z) = \text{step}(\mathbf{w}\mathbf{x})$.

La función step mas común en los perceptrones es la Heaviside, algunas veces

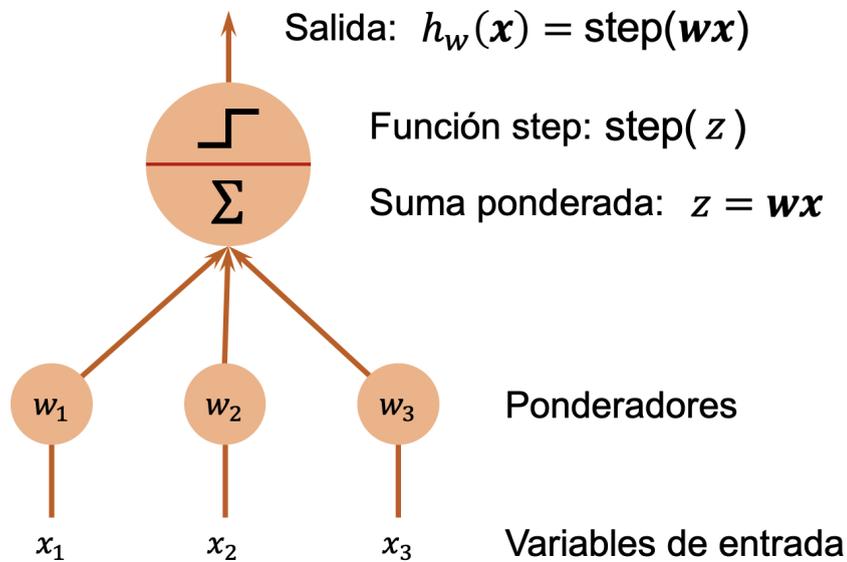


Figura 2.7: Unidad de umbral lineal (LTU, por sus siglas en inglés), (Crédito: Géron [30]).

también se usa la función signo, que se definen de la siguiente manera:

$$\text{heaviside}(z) = \begin{cases} 0 & \text{si } z < 0 \\ 1 & \text{si } z \geq 0 \end{cases}$$

$$\text{signo}(z) = \begin{cases} -1 & \text{si } z < 0 \\ 0 & \text{si } z = 0 \\ +1 & \text{si } z > 0 \end{cases}$$

Una sola LTU puede ser utilizada como un clasificador lineal binario. Calcula la combinación lineal de las entradas y si el resultado excede el umbral entonces genera la clase positiva y si no, genera la clase negativa justo como el clasificador SVM. Un perceptrón simplemente está compuesto por una capa de varios LTU con cada neurona conectada a todas las entradas. Esas conexiones se suelen representar usando neuronas de paso llamadas neuronas de entrada, éstas sólo emiten la salida de acuerdo al valor de entrada con el que son alimentadas, también se agrega una característica extra de sesgo ($x_0 = 1$). Esta característica se suele representar con una neurona especial

llamada neurona de sesgo, la cual solamente emite el valor 1 todo el tiempo.

Un perceptrón con dos entradas y tres salidas se representa en la figura 2.8. Éste puede clasificar ejemplos simultáneos de tres clases binarias lo cual constituye un clasificador multi-salida.

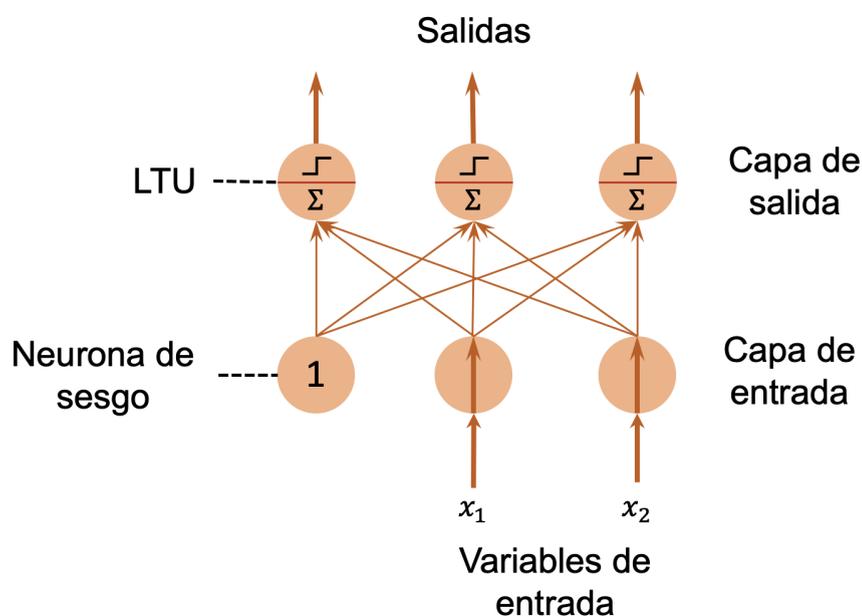


Figura 2.8: Diagrama de un perceptrón, (Crédito: Géron [30]).

La frontera de decisión de cada neurona de salida es lineal, por lo tanto, los perceptrones son incapaces de aprender de patrones complejos. Sin embargo, si las instancias de entrenamiento son linealmente separables, el algoritmo convergerá a una solución.

Perceptrón Multicapa

Algunas limitaciones de los perceptrones se pueden eliminar apilando múltiples perceptrones. La red neuronal artificial resultante es llamada perceptrón multicapa (MLP, por sus siglas en inglés de Multi Layer Perceptron). Un MLP se compone por una capa de entrada, una o más capas de LTUs, llamadas capas ocultas, y al final una capa de LTUs llamada capa de salida [30], en la figura 2.9 se muestra el esquema de un perceptrón multicapa. Cada capa, excepto la capa de salida, incluye una neurona de sesgo y está totalmente conectada a la siguiente capa. Cuando una red neuronal

artificial tiene dos o más capas ocultas, es llamada red neuronal profunda.

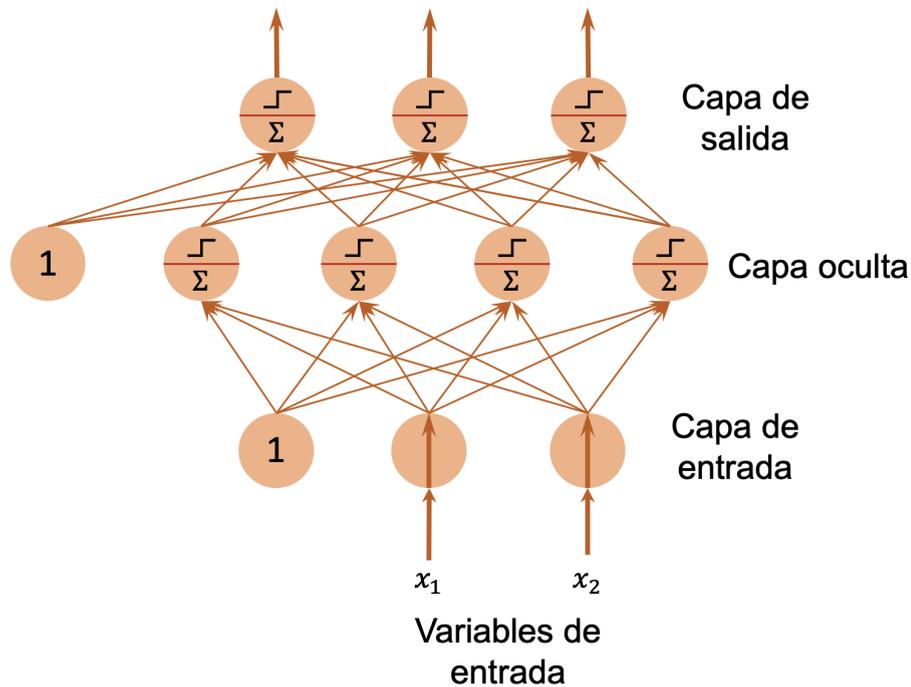


Figura 2.9: Diagrama de un perceptrón multicapa, (Crédito: Géron [30]).

Para cada instancia de entrenamiento, el algoritmo alimenta a la red y calcula la salida de cada neurona en cada capa consecutiva (esto es el paso hacia delante, justo como cuando se hacen predicciones). Entonces se mide el error de salida de la red (por ejemplo la diferencia entre la salida deseada y la salida real en este momento), con lo cual se calcula qué tanto cada neurona, de la última capa oculta, contribuye al error de salida por neurona. Entonces se procede a medir la contribución del error que proviene de cada neurona en la capa oculta anterior y así, hasta que el algoritmo alcanza la capa de entrada. Este paso inverso mide eficientemente el gradiente del error a través de todas las conexiones identificando el gradiente del error hacia atrás en la red. El último paso del algoritmo de propagación hacia atrás es el ajuste mediante gradiente descendiente en todas las ponderaciones de la red usando la medida del error calculada anteriormente. Es decir, para cada instancia de entrenamiento, el algoritmo de propagación hacia atrás primero realiza una predicción hacia delante, mide el error, y entonces va a través de toda la red en reversa para medir la contribución en el error de cada conexión. Finalmente se ajustan levemente los ponderadores de cada conexión para reducir el error.

2.3. Selección de Características

En esta sección explicaremos los fundamentos del proceso de selección de características utilizados en este trabajo. Es importante señalar que, en ambos conjuntos de datos de imágenes, el uso de características adicionales relacionadas con los análisis estadísticos, de textura y de forma mejoran significativamente los resultados en la tarea de clasificación LULC con respecto a solamente usar las bandas originales combinadas con índices espectrales. Debido al gran número de características generadas para cada conjunto de datos, consideramos relevante aplicar un método de selección de características.

El método usado en este estudio se denomina ReliefF a gran escala [22] (VLSReliefF), que es una variante del conocido método ReliefF [57,92] adaptado para manejar una gran cantidad de características. A su vez, el método ReliefF se basa en el método fundamental llamado Relief [56], por lo que iniciaremos la explicación de esta familia de métodos con este último.

2.3.1. Selección de características para dos clases

De acuerdo a Kira et al. y Robnik et al. [56, 78], el método Relief tiene como entrada un vector de atributos y la etiqueta de la clase para cada instancia de entrenamiento, como se muestra en el Algoritmo 1 (esquema mostrado mas adelante). Y como salida se obtiene un vector W con las estimaciones de calidad de cada atributo del vector de características. Se inicializan los valores del vector W en cero (línea 1). Posteriormente se selecciona aleatoriamente un ejemplo R_i , del conjunto de entrenamiento (línea 3). Lo cual se va a repetir m veces, según lo determine el usuario. Después se busca el vector de entrenamiento más cercano con la misma etiqueta de clase que el vector R_i , al cual se le llamará H (acierto), también se busca el ejemplo de entrenamiento más cercano con la etiqueta de entrenamiento distinta al que se le llamará M o desacierto (línea 4). Como indican en su trabajo Kira et al. y Robnik et al. [56, 78] ahora se actualizará la estimación de calidad para cada atributo individualmente, dependiendo de sus valores R_i , M y H (líneas 5 y 6). Si las instancias R_i y H tienen diferentes valores en el

atributo individual A , entonces dicho atributo separa dos instancias de la misma clase, lo cual no es deseable por lo que se decrementa la estimación de calidad de $W[A]$. Por el otro lado si las instancias R_i y M tienen diferentes valores en el atributo A entonces el atributo A separa dos instancias con diferente clase, lo cual es deseable, así que se incrementa la estimación de calidad $W[A]$.

La función $\text{diff}(A, I_1, I_2)$ calcula la diferencia entre los valores de un atributo A para dos instancias I_1 e I_2 . Para atributos nominales se definió como:

$$\text{diff}(A, I_1, I_2) = \begin{cases} 0 & \text{valor}(A, I_1) = \text{valor}(A, I_2) \\ 1 & \text{de otra manera} \end{cases} .$$

Y en el caso de que los atributos sean numéricos entonces se define como:

$$\text{diff}(A, I_1, I_2) = \frac{|\text{valor}(A, I_1) - \text{valor}(A, I_2)|}{\max(A) - \min(A)}$$

La función diff también se utiliza para calcular la distancia entre instancias. El algoritmo Relief está limitado a problemas con dos clases y depende de que los datos estén completos en todos sus valores, por lo que se desarrolló una extensión denominada ReliefF.

Algoritmo 1: Algoritmo Relief

Entrada: Para cada instancia de entrenamiento un vector valores de atributos y la etiqueta correspondiente

Salida: Vector W con las estimaciones de la calidad de los atributos

1 Establecer todos los pesos $W[A] := 0$

2 **para** $i:=1$ **hasta** m **hacer**

3 seleccionar aleatoriamente una instancia de entrenamiento R_i

4 encontrar el acierto H más cercano y el desacierto M más cercano

5 **para** $A:=1$ **hasta** a **hacer**

6 $W[A] := W[A] - \frac{\text{diff}(A, R_i, H)}{m} + \frac{\text{diff}(A, R_i, M)}{m}$

2.3.2. Selección de características multiclase

El algoritmo ReliefF (Relief-F) [58, 78] (ver Algoritmo 2) no se limita a problemas de dos clases y puede lidiar con datos ruidosos. Similar al algoritmo Relief, ReliefF selecciona una instancia R_i (línea 3), entonces se buscan k instancias de sus vecinos más cercanos que tengan la misma clase, llamados H_j aciertos más cercanos (línea 4). También se buscan los k vecinos más cercanos para cada una de las clases diferentes, llamados los $M_j(C)$ vecinos más cercanos (líneas 5 y 6). Ahora se actualiza la estimación de la calidad para cada atributo A dependiendo de los valores de R_i , aciertos H_j y desaciertos $M_j(C)$.

Algoritmo 2: Algoritmo ReliefF

Entrada: Para cada instancia de entrenamiento un vector valores de atributos y la etiqueta correspondiente

Salida: Vector W con las estimaciones de la calidad de los atributos

1 Establecer todos los pesos $W[A] := 0$

2 **para** $i:=1$ hasta m **hacer**

3 seleccionar aleatoriamente una instancia de entrenamiento R_i

4 encontrar los k aciertos H_j más cercanos

5 **para** cada clase $C \neq \text{clase}(R_i)$ **hacer**

6 de la clase C seleccionar los k desaciertos $M_j(C)$

7 **para** $A:=1$ hasta a **hacer**

8

$$W[A] := W[A] - \sum_{j=1}^k \frac{\text{diff}(A, R_i, H_j)}{mk} + \sum_{C \neq \text{clase}(R_i)} \frac{\frac{P(C)}{1-P(\text{clase}(R_i))} \sum_{j=1}^k \text{diff}(A, R_i, M_j(C))}{mk}$$

La fórmula con la que se actualizan las estimaciones de calidad $W[A]$ es similar a la de Relief en la línea 6 del Algoritmo 1. Con la excepción de que ahora se promedia la contribución de todos los aciertos y todos los desaciertos. La contribución de los desaciertos de cada clase se pondera con la estimación de la probabilidad a priori de cada clase $P(C)$ usando el conjunto de entrenamiento. Para lidiar con los datos incompletos la función diff se modificó. Los valores faltantes de los atributos

tos son tratados probabilísticamente. Se calcula la probabilidad de que dos instancias dadas tengan valores diferentes para un atributo dado condicionado sobre el valor de la clase. Si una instancia, por ejemplo I_1 , tiene un valor desconocido entonces: $\text{diff}(A, I_1, I_2) = 1 - P(\text{value}(A, I_2) | \text{clase}(I_1))$; si ambas instancias tienen valores desconocidos entonces: $\text{diff}(A, I_1, I_2) = 1 - \sum_V^{\#values(A)} (P(V | \text{clase}(I_1)) \times P(V | \text{clase}(I_2)))$. Las probabilidades condicionales se aproximan con frecuencias relativas a partir del conjunto de entrenamiento.

En la figura 2.10 se puede ver una ilustración del esquema de selección de vecinos más cercanos. El elemento seleccionado aleatoriamente está seleccionado con una estrella. En el caso de Relief, 1) del lado izquierdo, solamente se selecciona una instancia del espacio A-dimensional por cada tipo de vecino, es decir uno para la instancia que coincide con la clase de el ejemplo seleccionado aleatoriamente que está marcado con un punto y uno para la selección que no coincide con la instancia seleccionada la cual está marcada con una equis. Para ReliefF, 2) del lado derecho, el parámetro k es igual a 3, es decir se seleccionarán 3 vecinos coincidentes con la clase del seleccionado aleatoriamente y 3 vecinos para cada una de las clase no coincidentes.

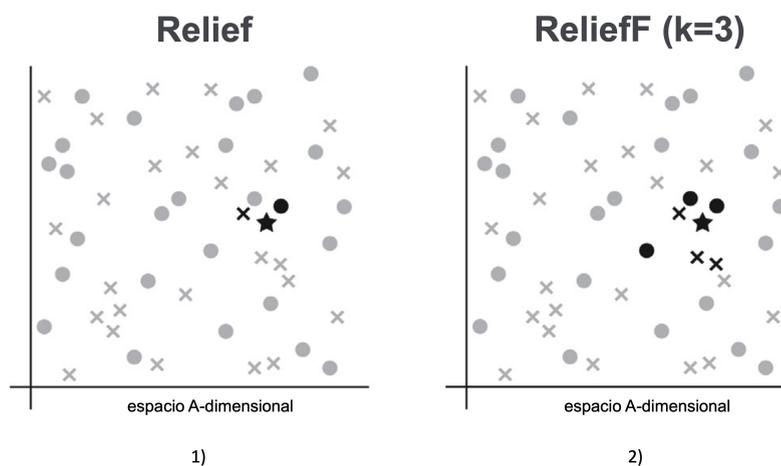


Figura 2.10: Comparación de esquemas de selección de vecinos más cercanos; 1) Selección para el algoritmo Relief, 2) Selección para ReliefF, (Crédito: Urbanowicz [91])

2.3.3. Selección de características para muy altas dimensiones

VLSReliefF (también conocido como ReliefF a gran escala, en inglés se conoce como: Very Large Scale ReliefF). Se sabe que las estimaciones de calidad W usando el

método ReliefF son más exactas cuando se aplican a subconjuntos pequeños de características [22], por lo que el principio detrás del algoritmo VLSReliefF (ver Algoritmo 3) es sencillo y consiste en aplicar ReliefF a subconjuntos S de atributos seleccionados aleatoriamente (línea 3). Cada subconjunto tiene un tamaño menor al total de atributos, cada uno de tamaño $N_s < N$. Después es posible combinar los resultados parciales de múltiples maneras. En éste algoritmo solamente se actualizan los pesos globales de cada atributo con el peso máximo local calculado en cualquier subconjunto (línea 13). De esta manera se aprovecha el poder selectivo del algoritmo ReliefF en subconjuntos pequeños de características combinando los resultados mediante la selección de las

mejores estimaciones de calidad en cada iteración.

Algoritmo 3: Algoritmo ReliefF

Entrada: Para cada instancia de entrenamiento hay un vector de valores de atributos y la etiqueta correspondiente

Salida: Vector W con las estimaciones de la calidad de los atributos

```

1 Establecer todos los pesos  $W[A] := 0$ 
2 para subconjunto:=1 hasta  $n$  Subconjuntos hacer
3      $S =$  Selección aleatoria de  $N_s$  atributos
4     ****ReliefF se aplica al subconjunto  $S$ 
5     para  $i:=1$  hasta  $m$  hacer
6         seleccionar aleatoriamente una instancia de entrenamiento  $R_i$ 
7         encontrar los  $k$  aciertos  $H_j$  mas cercanos
8         para cada clase  $C \neq \text{clase}(R_i)$  hacer
9             de la clase  $C$  seleccionar los  $k$  desaciertos  $M_j(C)$ 
10        para  $A:=1$  hasta  $a$  hacer
11            
$$W[A] := W[A] - \sum_{j=1}^k \frac{\text{diff}(A, R_i, H_j)}{mk} + \sum_{C \neq \text{clase}(R_i)} \frac{\frac{P(C)}{1-P(\text{clase}(R_i))} \sum_{j=1}^k \text{diff}(A, R_i, M_j(C))}{mk}$$

12        ****Concluye ReliefF
13     $W[A]_S$  se actualiza si la estimación máxima local es mayor que la
        estimación actual de cada atributo

```

2.4. Resumen

En este capítulo se presentó un resumen general de los conceptos básicos de percepción remota y aprendizaje automático.

En el caso de la percepción remota se revisó el concepto de espectro electromagnético y su influencia en el diseño de sensores satelitales pasivos, dónde se establecen rangos de observación de dicho espectro, a los que se denomina bandas es-

pectrales. A partir de la cuantificación de la energía recibida en esas regiones de observación, es posible obtener información de cómo interactúa la radiación solar con los objetos en la superficie terrestre, lo cual es el fundamento de la percepción remota y está relacionado con el concepto de resolución espectral que tiene que ver con la cantidad de subdivisiones independientes (bandas) que puede registrar el satélite de observación de la tierra. Por último también se revisaron las diferentes resoluciones que puede tener un satélite de percepción remota, como lo son:

1) la resolución temporal, que tiene que ver con el tiempo que tarda un satélite en concluir una cobertura completa de la tierra.

2) La resolución espacial, que está relacionada con el máximo detalle que puede captar un satélite de la superficie terrestre, es decir, el tamaño mínimo que cada píxel representa de la superficie terrestre.

3) La resolución radiométrica, que tiene que ver con el tamaño de la representación binaria usada para almacenar la información registrada por los sensores.

Todo lo anterior es integrado para generar diferentes tipos de productos de acuerdo al nivel de procesamiento que sufren a la hora de generarse.

En el caso de este trabajo utilizamos imágenes Landsat con el mayor nivel de procesamiento disponible hasta el momento en que se realizó el trabajo denominado *Analysis Ready Data* (ARD).

En la segunda parte de este capítulo, se revisaron los algoritmos fundamentales de aprendizaje automático usados en este documento. La primera familia de algoritmos revisada fue la de los métodos basados en árboles, ahí se mostraron los elementos básicos que componen a un algoritmo de clasificación llamado árbol de decisión, el cual tiene la ventaja de ser muy fácil de interpretar y fácil de entrenar, aunque no poseen grandes capacidades predictivas comparadas con otros modelos. Sin embargo, usando los principios básicos de los árboles de decisión, es posible incorporarle mejoras aplicando técnicas de reducción de varianza basadas en el re-muestreo como el *bootstrapping*, construyendo múltiples árboles de decisión utilizando subconjuntos de los datos de entrenamiento obtenidos a partir de re-muestreo; así como también, se-

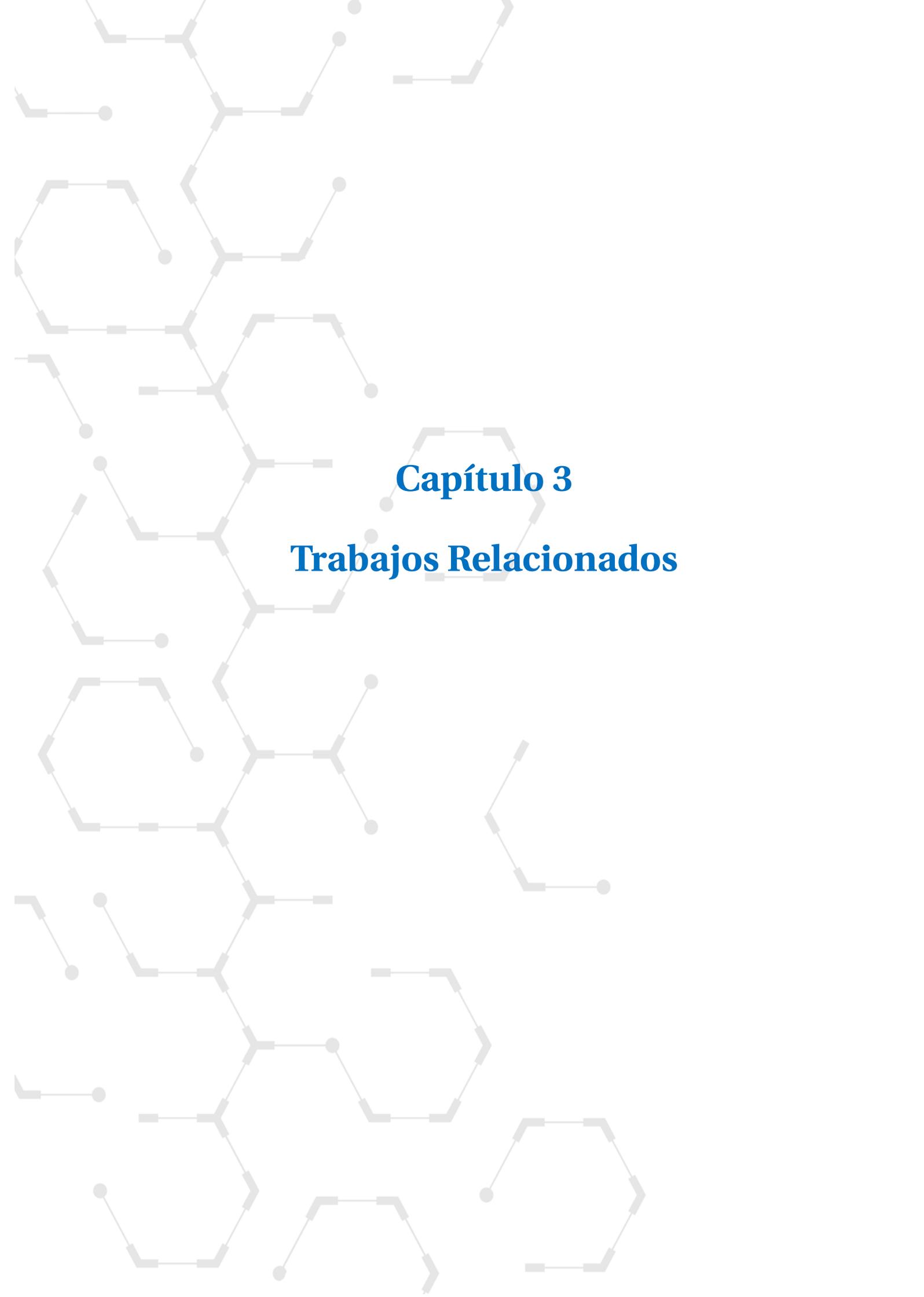
leccionando aleatoriamente subconjuntos de características para realizar un ensamble de múltiples árboles aleatorios, mejorando significativamente las capacidades predictivas del algoritmo a costa de la interpretabilidad. Todo esto, debido a que ya no es posible representarlos como un solo árbol de decisión. A este tipo de algoritmos se les conoce como Bosque Aleatorio.

Yendo más allá en el proceso de aleatorización, se genera un nuevo tipo de clasificador basado en árboles, el cual selecciona aleatoriamente los umbrales para generar la segmentación del espacio de características, en lugar de buscar el umbral óptimo con las características disponibles para cada árbol que se construye. Aumentando así significativamente la velocidad del entrenamiento sin afectar la capacidad predictiva del algoritmo.

También se revisaron los principios del algoritmo de clasificación llamado Support Vector Machines (SVM). Éste busca identificar la frontera de decisión que separa dos clases: utilizando principios geométricos, identificando vectores de soporte en el conjunto de entrenamiento y generando la frontera de decisión dando el máximo margen posible a las dos clases involucradas. Lo anterior, como una estrategia de generalización que permitirá mantener sin cambios al modelo, incluso si llegan nuevos datos de entrenamiento dentro de los márgenes establecidos. SVM es un método eficiente y frecuentemente logra buenos resultados en comparación con otros algoritmos de clasificación.

Finalmente, se revisaron las Artificial Neural Networks empezando por la neurona básica denominada unidad de umbral lineal, la cual se ensambla en una arquitectura muy sencilla llamada perceptrón, que solamente puede lidiar con problemas linealmente separables. Sin embargo, para superar algunas de sus limitaciones, se apilan múltiples capas de estos perceptrones generando un perceptrón multicapa, el cual es entrenado utilizando una técnica de medición del error hacia adelante y propagación hacia atrás, con el que se ajustan los parámetros del modelo de forma iterativa, logrando con esto, un modelo de clasificación supervisada muy efectivo en el proceso de clasificación en comparación con otros métodos.

Por último, en la tercera parte se explicaron algunos algoritmos de selección de características dónde se realizan evaluaciones de la capacidad de separación de las clases que tiene cada atributo del conjunto de características, hasta llegar al algoritmo VLSReliefF usado en este trabajo.



Capítulo 3
Trabajos Relacionados

Capítulo 3 Trabajos Relacionados

3.1. Problemas de aprendizaje Supervisado en Percepción Remota

Se han propuesto varios esfuerzos para abordar el problema de clasificación LULC, por ejemplo, en Xiong et al. [100], se lleva a cabo el análisis de grandes volúmenes de imágenes de satélite Landsat y Sentinel-2 para construir dos compuestos continentales de África con píxeles a una resolución de 30 metros para detectar las zonas de cultivo de todo el continente. A partir de los dos compuestos continentales, correspondientes a dos periodos subanuales, las características de cada uno fueron los promedios de cuatro canales espectrales: azul, verde, rojo e infrarrojo cercano (NIR), y el índice de vegetación de diferencia normalizada (NDVI). En particular, el NDVI se calcula como la diferencia normalizada entre la reflectancia en el infrarrojo cercano y la reflectancia en el rojo, dividido por la suma de ambas. El NDVI toma valores entre -1 y 1, siendo los valores más altos indicativos de una mayor densidad de vegetación. Adicionalmente, se usó la pendiente proveniente del modelo digital de elevación de la misión topográfica shuttle radar (SRTM) que existe para la mayor parte del planeta. Un modelo digital de elevación (MDE) es una representación en 3D de la superficie terrestre, que muestra la elevación del terreno en cada punto. Como resultado, se usaron 11 características y se usó un clasificador de Random Forest (RF) para detectar las clases de cobertura del suelo de cultivo y sin cultivo, logrando una exactitud general de 94.5%.

Un aspecto destacable de este estudio, es la combinación de imágenes multiespectrales Sentinel-2 re-muestreadas a 30 metros para construir los dos continuos continentales para África, sin embargo, no les fue posible lograr un cubrimiento completo de todo el continente, debido a la presencia constante de nubes en ciertas regiones, por lo que se completó la información con píxeles provenientes de imágenes Landsat-8 del mismo periodo con la misma resolución espacial de 30 metros.

El trabajo presentado en Jamsran et al. [49], se llevó a cabo la clasificación de cobertura de suelo y el cambio en la cobertura de suelo de la región ecológica de la cuenca del Lago Uvs de Mongolia a partir de imágenes Landsat de los años de 1995 a 2015. Las clases identificadas fueron bosque de coníferas, bosque latifoliado, pradera, tierras desoladas, tierra de roca desnuda, cuerpos de agua, arena y glaciar. En el trabajo, se usaron como características las bandas Landsat a 30 metros de resolución azul, verde, rojo, NIR, SWIR1, SWIR2 y lograron un 92 % de exactitud global usando el clasificador Support Vector Machines (SVM). La región estudiada es relativamente pequeña debido a que solamente tiene un área de 150 km × 150 km, los cuales se pueden obtener de una sola imagen Landsat. Sin embargo, se realizó el estudio con imágenes que van desde 1995 hasta 2015, con lo que logran estudiar la evolución de la cobertura de suelo a lo largo de 20 años.

El resultado de la clasificación de la cobertura de suelo mediante técnicas de Machine Learning, permitió identificar las zonas de mayor degradación de bosques debida a la actividad humana y al cambio climático. El estudio de 20 años de historia en la cobertura de suelo correspondiente al estudio se dio gracias a que el acervo Landsat tiene almacenadas varias décadas imágenes de toda la Tierra.

En el trabajo de Shen et al. [86], se utilizan múltiples clasificadores combinados mediante votaciones ponderadas para identificar nueve tipos de cobertura de suelo (bosque de eucaliptos, bosque de manglar, bosque de coníferas, bosque latifoliado, bosque mixto, campo de caña de azúcar, campo de arroz, cuerpos de agua y áreas urbanas) en imágenes Landsat de la región vecina al Golfo de Beibu, situado en el noroeste del mar de China meridional, logrando una exactitud global de 93.05 %. En éste trabajo se usaron como características los valores de las bandas de aerosol costero, azul, verde, rojo, NIR, SWIR1, SWIR2 y cirrus, todas a 30 metros de resolución, así como la banda pancromática con 15 metros de resolución. Se explora el potencial de los ensambles de clasificadores, compuestos por clasificadores más simples para realizar tareas de análisis de imágenes Landsat para detectar clases de cobertura de suelo.

En el trabajo de Picoli et al. [76], se utilizan series de tiempo de imágenes de los índices de vegetación del producto, derivado del sensor MODIS llamado MOD13Q1, el

cual se genera cada 16 días con píxeles a 250 m de resolución espacial; el periodo estudiado está entre los años 2001 y 2016. La región de análisis corresponde al estado de Mato Grosso en Brasil. En el estudio se identificaron nueve clases de cobertura y uso de suelo (bosque, sabana tropical, pastizales y los cultivos de soya, algodón, soya-almorran, soya-maíz, soya-mijo y soya-girasol). Las características usadas fueron cuatro bandas del producto MOD13Q1, las cuales son: NDVI, EVI, NIR e infrarrojo medio (MIR), con las que lograron una exactitud global de 93 % con un clasificador SVM. Los autores utilizan series de tiempo anuales a nivel píxel, usando toda la información de las cuatro características a lo largo de cada año analizado, para construir un perfil temporal de cada cultivo; con lo que logran caracterizar el comportamiento de cada cultivo a lo largo del año, llevar a cabo el entrenamiento del clasificador SVM.

En el estudio presentado en Kussul et al. [61] se lleva a cabo el uso de imágenes multiespectrales de los satélites Landsat y Sentinel-2 e imágenes de radar del satélite Sentinel-1A. Se estudiaron los años, 2013, 2014 y 2015 del distrito de Bilotserkivskiy de la región de Kyiv en Ucrania. Se identificaron 13 clases (estructuras artificiales, trigo, canola, maíz, cultivos de primavera, remolacha, girasol, soya, otros cereales, bosque, pradera, tierras desoladas y cuerpos de agua). Se usaron como características las polarizaciones radar VV y VH remuestreadas a 30 metros para ser combinadas con las bandas azul, verde, rojo, NIR, SWIR1 y SWIR2 del satélite Landsat 8. También se usaron las bandas azul, verde, rojo y NIR a 10 metros de resolución provenientes del satélite Sentinel-2. Usando el promedio de múltiples perceptrones multicapa se lograron las exactitudes globales de 85.3 % para 2013, 90.1 % para 2014 y 92.4 % para 2015.

En este estudio los autores usan información proveniente de satélites ópticos (Landsat y Sentinel-2), sin embargo, en las épocas del año donde no pueden obtener información debido a la presencia casi permanente de nubes, ellos la complementaron con imágenes provenientes del satélite Sentinel-1, el cual atraviesa las nubes por basarse en la tecnología del radar, y con lo cual fusionaron múltiples tipos de imágenes para completar la información en todo el periodo de tiempo.

En el trabajo Phiri et al. [18], se llevó a cabo una revisión de múltiples estudios realizados sobre la clasificación de cobertura de suelo usando imágenes Landsat 4, 5,

7 y 8, a partir de lo cual identificó que la técnica de análisis de imágenes basadas en objetos (OBIA, de sus siglas en inglés) la cual trabaja con segmentaciones homogéneas de píxeles, logra los mejores niveles de exactitud global que van desde 80% hasta 97%. Las clases de cobertura de suelo estudiadas fueron agrupadas como: áreas urbanas, plantación forestal, bosque natural y zona agrícola. Además de los valores espectrales, la técnica de clasificación OBIA también incorpora información proveniente de la textura, forma y posición de las imágenes debido a que trabaja con agrupaciones de píxeles.

Los autores identifican que la estrategia OBIA de clasificación se encuentra como una tendencia creciente en el estado del arte, demostrando exactitudes notables en comparación a otras estrategias. Por lo que sugiere considerar las fortalezas y debilidades del método y continuar realizando estudios que realicen comparaciones con los algoritmos tradicionales.

En la propuesta presentada en Teluguntla et al. [88], se construyeron seis compuestos subanuales con todas las imágenes Landsat disponibles para Australia, y cuatro compuestos subanuales de las imágenes de China, para los años 2013 a 2015 en ambos casos. Para el conjunto de datos de cada país, se utilizaron ocho características espectrales: azul, verde, rojo, NIR, infrarrojo de onda corta 1 (SWIR1), infrarrojo de onda corta 2 (SWIR2), infrarrojo térmico (TIR) y el índice NDVI; lo que resulta en 48 características para Australia y 32 para China. Luego, se aplicó un clasificador Random Forest (RF), logrando una exactitud global del 97.6% para las imágenes de Australia y del 94% para China para la clasificación de las clases cultivos, sin cultivos y pastizales.

Los autores utilizan la plataforma pública de Big Data en la nube llamada Google Earth Engine (GEE), con la cual logran analizar grandes extensiones territoriales con imágenes de resolución media como lo son las Landsat de 30 metros por pixel. El volumen de información procesado en el trabajo asciende a la escala de petabytes según refieren los autores, por lo que la solución propuesta es un cambio de paradigma a considerar para futuros trabajos.

Los autores Feng et al. [23] proponen el cálculo de 145 características a partir

de imágenes Landsat agrupadas en cuatro períodos subanuales sobre África desde el año 2015. Con base en estas imágenes, se utilizaron siete bandas espectrales (aerosol costero, azul, verde, rojo, NIR, SWIR1 y SWIR2) y se extrajeron siete índices espectrales, NDVI, índice de vegetación mejorado (EVI), índice de vegetación ajustado al suelo (SAVI), Índice de vegetación resistente a la atmósfera (ARVI), índice de agua de diferencia normalizada modificada (MNDWI), ver Tabla 1.1. Además, se calcularon características según un conjunto anual de los índices de vegetación MODIS, MOD13Q1, de los cuales se extrajo el índice NDVI. A partir de los datos MODIS, se calculó un conjunto de seis valores estadísticos: media, valor máximo, valor mínimo, desviación estándar, varianza y mediana.

El modelo de elevación digital (MDE), los datos climáticos y las luces nocturnas de las observaciones satelitales de la serie de radiómetros de imágenes visibles mediante infrarrojo (VIIRS) en su banda día / noche (DNB), también se utilizaron como características. Finalmente, se logró una exactitud general de 75.8% para la clasificación de 10 clases (tierras de cultivo, bosques, praderas, matorrales, humedales, agua, áreas impermeables, tierras desoladas, nieve y hielo, nubes) utilizando un clasificador de RF. En este trabajo es destacable la integración de múltiples fuentes para la elaboración de un mapa de cobertura de suelo para el continente africano.

Heydari y otros [85] estudian una muestra de 26 imágenes Landsat de la región continental de Estados Unidos para el año 2011, donde se identificaron 11 clases (cuerpo de agua, área urbana, perturbada mecánicamente, tierras desoladas, minería, bosques, pastizales o matorrales, zona agrícola, humedal, perturbada no mecánicamente y hielo o nieve). Como características, se usaron directamente 6 bandas (azul, verde, rojo, NIR, SWIR1 y SWIR2) a 30 metros de resolución espacial, logrando una exactitud global de 84.7%, usando un clasificador SVM.

Los autores llevan a cabo la evaluación de seis métodos de clasificación: Naïve Bayes (NB), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Bootstrap Aggregation Ensemble of Decision Trees (BagTE), Artificial Neural Network (ANN) y una Deep Neural Network (DNN), siendo SVM el clasificador que mostró mejor desempeño global y cómo segundo mejor clasificador resultó ser KNN en los experimentos

realizados con diversas estrategias de muestreo. Los autores advierten que es importante la realización de búsquedas extensas en sus espacios de parámetros para lograr buenos resultados, así como contar con muestras de entrenamiento equilibradas para cada clase.

En el trabajo Cai, Guan et al. [12] proponen el uso de series temporales de interpolaciones diarias de valores espectrales (azul, verde, rojo, NIR, SWIR1, SWIR2) y varios índices de vegetación como NDVI, índice de vegetación de clorofila verde (GCVI), EVI y el índice de agua superficial (LSWI) el cual se calcula a partir de las bandas espectrales del infrarrojo cercano (NIR) y el infrarrojo de onda corta 1 (SWIR1), y su fórmula se obtiene al restar el valor de SWIR1 del valor de NIR y dividir el resultado entre la suma del valor de NIR y SWIR1. Luego, con una DNN como clasificador, se obtuvo el resultado de una exactitud general del 97% en las clases de maíz y soya. En este trabajo, los autores realizan un perfilado de cada cultivo mediante las series de tiempo de cada banda espectral, así como de los índices de vegetación elegidos, con los cuales identifican que la combinación de bandas espectrales más predicativas resultaron ser SWIR-1 y verde. Esto es notable debido a que el autor indica que la banda SWIR-1 no suele usarse en detección de cultivos por lo que, dicho hallazgo es relevante. También indicaron en el estudio que los índices de vegetación no superaron los resultados de exactitud logrados directamente por las bandas espectrales. Por último se establece que futuras mejoras podrían consistir en incorporar características de textura a partir de series de tiempo provenientes de información de múltiples sensores como MODIS y Landsat.

Las imágenes de Landsat de seis años (1995, 2000, 2005, 2010, 2015 y 2017) se utilizan en Teffera et al. [87]. Aquí, los autores se centran en las seis bandas espectrales (azul, verde, rojo, NIR, SWIR1 y SWIR2) y ocho índices espectrales: índice urbano (UI), SAVI, MNDWI, NDVI, índice de agua de diferencia normalizada (NDWI), índice de acumulación de diferencia normalizada (NDBI), índice acoplado urbano vegetativo (CVUI) e índice de suelo desnudo (BI), ver Tabla 1.1. Se clasificaron seis clases diferentes: tierras áridas o desarrolladas, bosques, tierras agrícolas de riego, tierras agrícolas de temporal, cuerpos de agua y tierras boscosas o con arbustos. Además, se utilizaron

combinaciones de tres índices de vegetación para hacer compuestos de falso color para generar una metodología de mapeo de las seis clases consideradas a lo largo de los años estudiados. Con esta metodología, se alcanzó una exactitud general de 91.8% con el clasificador de máxima verosimilitud. En este estudio llevaron a cabo la aplicación de la técnica de Análisis de Componentes Principales (PCA, por sus siglas en inglés) para reducir las dimensiones espectrales, en el caso de los índices de espectrales usados, los autores crearon imágenes con combinaciones de éstos, revelando su potencial para discriminar fácilmente diversas características del terreno como agua, vegetación y suelo desnudo. Lo cual es particularmente útil cuando se monitorea la dinámica de la cobertura de suelo a lo largo del tiempo.

En el trabajo de Phalke y otros [75], los autores utilizaron series temporales de imágenes Landsat para clasificar una gran región de Europa, Medio Oriente y África del Norte en las clases de cultivo y sin cultivo. Calcularon el índice espectral EVI del que extrajeron siete características estadísticas (media, desviación estándar, máximo, mínimo, varianza, rango y frecuencia de los valores del índice para cada píxel) de la serie de tiempo, además de la elevación y la pendiente del modelo digital de elevación. El resultado fue una exactitud general del 74% con el clasificador Análisis Discriminante Lineal (LDA, por sus siglas en inglés). Los autores prueban la idea de generalización en la generación de mapas de cultivos (cultivo, sin cultivo) logrando verificar que se producen resultados aceptables usando muestras de entrenamiento de la misma época pero de diversas regiones, que el modelo entrenado puede ser aplicado para clasificar imágenes de todas las regiones; abriendo la posibilidad de usar modelos de clasificación en áreas continentales, sin la necesidad de que se realicen múltiples entrenamientos.

Con imágenes Landsat 8 de la región de Cape Winelands, Sudáfrica, se realizó el trabajo presentado por Gilbertson et al. [50] el cual se enfoca en la identificación de siete clases agrícolas (alfalfa, uva, canola, lupino, trigo, aceituna y pastizales) dónde se utilizan 41 características para realizar la clasificación supervisada. Se calculó la media y la desviación estándar de siete bandas (azul, verde, rojo, NIR, SWIR1, SWIR2 y pancromática), a partir de dichas bandas se generaron nueve índices espectrales: ARVI, EVI, índice de cobertura verde (GCI), índice de vegetación de diferencia normalizada verde

(GNDVI), índice de verdor (GI), NDVI, índice de la razón rojo verde (RGRI), SAVI, índice de razón simple (SR), ver Tabla 1.1, cuatro estadísticas derivadas de la matriz de co-ocurrencia de nivel de gris (GLCM) (Contraste, Correlación, Entropía y Homogeneidad) y tres estadísticas derivadas del vector de diferencias de nivel de gris (GLDV) (Contraste, Entropía y Media). También se usaron cuatro Componentes Principales derivadas de las medias de las bandas espectrales y seis componentes de la transformación tasseled cap (brillo, verdor, humedad, transformación 4, transformación 5 y transformación 6), con las que logró una exactitud global de 95.5% con un clasificador SVM. Además, se prueban múltiples clasificadores, Decision Trees (DT), K-Nearest Neighbors (KNN), RF y SVM, para detectar diferentes tipos de cultivo mostrando el valor que tiene aplicar dichos modelos a la tarea de detección de cultivos. También se evaluó el impacto positivo en la detección de cultivos que tiene el proceso de aumento de resolución mediante la fusión de las bandas multiespectrales con la banda pancromática de Landsat de mayor resolución.

Usando imágenes Landsat, en el trabajo presentado en Rodríguez-Galiano et al. [93], se lleva a cabo la clasificación de 14 categorías de cobertura de suelo (zona urbana, bosque de álamos, bosque de coníferas, invernaderos, matorrales, surcos de olivos, pastizales, bosque de robles, cultivos de riego, cultivos leñosos de riego, cultivos herbáceos de temporal, tierras desoladas, cultivos tropicales y cuerpos de agua) en Granada al sur de España, explorando las capacidades de diversos algoritmos de clasificación midiendo su tolerancia a ruido y baja cantidad de datos disponibles. Logrando así una exactitud global de 92% con los algoritmos RF y SVM. Las características usadas fueron siete bandas espectrales Landsat (azul, verde, rojo, NIR, SWIR1, SWIR2 e infrarrojo térmico) adicionalmente seis componentes de la transformación tasseled cap (brillo, verdor, humedad, transformación 4, transformación 5 y transformación 6). Por último 3 variables topográficas (elevación, pendiente y aspecto).

Los autores compararon cuatro clasificadores (DT, RF, ANN, SVM) en la tarea de clasificación de cobertura de suelo, logrando los mejores resultados con los clasificadores RF y SVM. Los autores también probaron la robustez de los algoritmos agregando ruido verificando que RF y SVM siguen siendo los mejores aún con presencia de ruido.

Por último también probaron el desempeño de los algoritmos con menores cantidades de datos de entrenamiento, siendo SVM el más robusto seguido por RF.

En el trabajo presentado en Luciano et al. [21], se identifica la presencia o ausencia del cultivo de caña de azúcar en el estado de Sao Pablo, Brasil. Se usaron imágenes Landsat del 2008 al 2016 de las cuales se usaron 10 características, donde cuatro fueron bandas espectrales (rojo, NIR, SWIR1 y SWIR2); seis índices espectrales: NDVI, EVI, SAVI, índice de humedad de diferencia normalizada (NDMI) y dos índices de agua de diferencia normalizada (NDWI1, NDWI2). Dónde NDWI1 se calcula restando el valor de reflectancia en la banda verde del valor de reflectancia en la banda del infrarrojo cercano (NIR), y dividiendo esta diferencia entre la suma de los dos valores y NDWI2 se calcula restando la banda roja de la banda infrarroja de onda corta 1 (SWIR1) y dividiéndola por la suma de las mismas dos bandas. Se logró una exactitud global del 96% para los años analizados usando el algoritmo de RF.

Los autores exploran las capacidades de generalización en espacio y tiempo para el problema de detección de caña de azúcar, resaltando la necesidad de construir conjuntos de datos consistentes a lo largo de tiempo para que puedan ser usados sistemáticamente en nuevas clasificaciones sin necesidad de otros entrenamientos, en regiones donde no se encuentren datos de campo. También establecen que la metodología propuesta puede ser aplicada a otros cultivos anuales.

Con imágenes Landsat de la República de Zambia en Africa en el trabajo de Azari y otros [24], se logro una exactitud global de 89%, usando un clasificador Random Forest (RF). Las clases identificadas fueron cultivos de temporal, cultivos de riego, vegetación natural de dosel abierto, vegetación natural del dosel cerrado, vegetación natural de pantano y zonas urbanas. Las características usadas fueron seis bandas (azul, verde, rojo, NIR, SWIR1, SWIR2) y dos índices (NDVI y GCVI) a partir de los cuales se generaron resúmenes estadísticos (los cuantiles 0.1, 0.25, 0.5, 0.75 y 0.9) por píxel de las imágenes Landsat disponibles para Zambia durante 2012 a 2015. Llevaron a cabo la creación de compuestos sub- anuales con la tecnología de GEE y generaron mapas de cobertura de suelo con una calidad superior a los existentes para la región de estudio, mostrando el potencial de las plataformas en la nube para el análisis de grandes

volúmenes de información.

Shaharum y otros [84], llevan a cabo la clasificación de cobertura de suelo de la reserva natural de vida salvaje Krau, Malasia, usando imágenes Landsat del año 2016. Se identificaron cinco clases (bosque denso, agricultura o bosque no denso, áreas urbanas, suelo desolado y cuerpos de agua). Para generar las características, se realizó un proceso de refinado pancromático dando como resultado seis bandas a 15 metros de resolución (aerosol costero, rojo, verde, azul, NIR, SWIR1, SWIR2). Se logró una exactitud global de 98.22% usando una Artificial Neural Network (ANN). Los autores compararon tres algoritmos de clasificación: Spectral Angle Mapper (SAM), SVM y ANN, fusionando la banda pancromática con las bandas multiespectrales para mejorar la resolución de las imágenes Landsat, pasando de 30 metros de resolución a 15 metros. Como resultado de la clasificación, generaron un mapa de cobertura y uso de suelo con el que podrán llevar a cabo el monitoreo de recursos de la reserva natural del estudio.

Para monitorear los manglares en la costa sureste de China e islas circundantes, en el trabajo de Hu y otros [66] se llevó a cabo el estudio de las imágenes Landsat disponibles para esa región desde 1986 a 2017. Como características del estudio se usaron seis bandas espectrales (azul, verde, rojo, NIR, SWIR1 y SWIR2) y a partir de éstas se calcularon tres índices espectrales (NDVI, NDWI y MNDWI), de dichas características se calcularon cinco cuantiles (0.1, 0.25, 0.5, 0.75 y 0.9) a lo largo de cada periodo estudiado. Adicionalmente se usaron los datos de altitud de un modelo digital de elevación para calcular tres características (elevación, pendiente y aspecto) y dos características geográficas provenientes de la rejilla Landsat (path y row). Se logró una exactitud de productor, la cual se refiere a los píxeles correctamente clasificados por clase entre el total de píxeles de cada clase, de 94.23% y una exactitud de usuario, la cual se refiere a los píxeles clasificados correctamente por clase entre los píxeles clasificados por clase, de 96.75% usando un clasificador de Random Forest (RF). Los autores emplearon perfiles temporales de bandas espectrales e índices en periodos multianuales, con lo que identificaron diferencias significativas entre el manglar y otros tipos de cobertura de suelo. Lograron identificar que el método propuesto basado en series de tiempo de imágenes, supera a los métodos que solamente utilizan una sola fecha.

En el trabajo de Gebhardt et al. [28] se propuso un enfoque de clasificación de series temporales de imágenes de Landsat de varios años en México. Aquí, los autores propusieron una metodología para clasificar 12 clases de cobertura del suelo (bosque de coníferas templado, bosque caducifolio templado, bosque mixto templado, bosque siempre verde tropical, bosque caducifolio tropical, matorral, vegetación de humedales, agricultura, pastizales, cuerpo de agua, tierra estéril y zona urbana). Para cada píxel, se calcularon 10 índices espectrales y seis componentes de la transformación tasseled cap. Luego, las series temporales de diez índices se resumieron en cinco estadísticas temporales (valor mínimo, valor máximo, rango, promedio y desviación estándar) generando 10 imágenes, una para cada índice con cinco dimensiones resultantes de cada estadística calculada. Más tarde, las imágenes se segmentaron generando cuatro resúmenes estadísticos (mínimo, máximo, promedio y desviación estándar). Como resultado, se generaron 200 características y se utilizaron como entrada de un método PCA. Del cual, se seleccionaron los componentes más significativos, que explican el 95% de la varianza. Además, se incluyeron el resumen estadístico (mínimo, máximo, promedio y desviación estándar) de la pendiente, y la información del modelo de elevación digital. Para todas estas características se logró una exactitud general del 71% para el año 2000 utilizando el clasificador de árbol de decisión supervisado llamado C5.0.

El conjunto de datos EuroSAT Sentinel-2 recientemente publicado por Helber y otros [36], se ha utilizado en múltiples estudios, el conjunto de datos se publicó con un punto de referencia del 98.57% de exactitud general utilizando una Convolutional Neural Network (CNN) ResNet-50 afinada. Los trabajos de Kroupi et al. [59], Roy et al. [82], Cavallaro et al. [14] y Senecal et al. [83] utilizan dicha base de datos, obteniendo los siguientes resultados: En el trabajo de Kroupi et al. [59] se reportó una exactitud de 87.6% usando un proceso basado en una CNN. En el trabajo de Roy y otros [82] se logró el 93.2% de exactitud general con una red generativa antagónica semi-supervisada. En Cavallaro et al. [14] se informó un 93.6% de exactitud general con una arquitectura de red convolucional basada en nodos de aprendizaje SVM (CSVM). En Senecal et al. [83] se logró un 96.6% con CNN para la clasificación de imágenes multiespectrales. En

esta tesis, como se verá más adelante, se utilizó este conjunto de datos para la parte experimental.

Buscando qué tipo de características han obtenido mejores resultados en otros trabajos, Hossain y otros [39] busca subconjuntos de características aplicando una medida normalizada de información mútua (MI) como un mecanismo de clasificación para cada una de las características. Los autores usan las bandas de salida del sensor directamente para encontrar la mejor combinación de características para lograr el mejor rendimiento en exactitud, aunque no informa qué características fueron las mejores para la tarea.

En Georganos y otros [29], se realizó la evaluación de 169 características derivadas del análisis estadístico, índices espectrales y textura del GLCM. En el estudio, hasta un 23 % de las características más significativas se identificaron mediante varios métodos de selección. Posteriormente, los grupos de características se probaron con varios métodos de clasificación buscando el mejor resultado de precisión. El autor exploró el potencial de los métodos de selección de características a partir del análisis estadístico de índices espectrales y de textura basados en GLCM e identificó un método de clasificación apropiado evaluándolo contra las múltiples selecciones resultantes.

En el trabajo de Zhou et al. [105], se propone un método de selección de características, que integra el método ReliefF con SVM para la identificación de edificios en imágenes de alta resolución. El autor realiza la selección de características de un conjunto de 85 que se agrupan en estadísticas basadas en índices espectrales, geométricos basados en la forma de los edificios, texturas basadas en GLCM, índices de sombra y características geográficas. Realiza la selección del 10 % más significativo, también explora el impacto en los tiempos de procesamiento cuando se utilizan todas las funciones frente a una selección menor. La selección de características importantes da como resultado una mejora en la precisión y los tiempos de procesamiento.

En el trabajo de Wang, Zhang et al. [95] se propone una variante del algoritmo ReliefF aplicado a 33 características de imágenes de alta resolución distribuidas en grupos de formas y texturas espectrales basadas en GLCM. Como resultado, se obtiene una

selección de siete características de los tres grupos: forma, textura e índice espectral. Si bien el número de características es muy pequeño en comparación con el realizado en el presente estudio, podemos observar que la combinación de diferentes tipos de características contribuye al proceso de clasificación.

3.2. Resumen

En la Tabla 3.1 se presenta un resumen de los principales aspectos de los trabajos mencionados. En la tabla se puede observar que el número de clases va desde dos hasta 14 con un promedio de siete en los trabajos mostrados.

También se observa que los métodos más usados son SVM, RF y diversas variantes de ANN.

En el caso de las características, prevalece el uso de las bandas espectrales directamente combinadas con índices espectrales.

Año	Referencia	Región	Sensor	Imágenes	Modelo	Clases	Características		O.A.
							Total	Tipos	
2018	[36]	Europa	Sentinel 2	27,000	Res-Net-50	10	3	CNN	98.57 %
2018	[59]	Europa	Sentinel 2	27,000	DCNN	10	4	CNN	87.6 %
2018	[82]	Europa	Sentinel 2	27,000	SF-GAN	10	3	CNN	93.2 %
2018	[14]	Europa	Sentinel 2	27,000	MCSVM	10	3	CSVM	93.6 %
2018	[83]	Europa	Sentinel 2	27,000	CNN	10	10	CNN	96.6 %
2018	[88]	Australia, China	Landsat	6, 4 Compuestos Nacionales	RF	3, 2	48, 32	Espectrales	94 %, 94 %
2014	[28]	México	Landsat	3600	C5.0	12	200	Espectrales, Temporales, Topográficas,	71 %
2017	[100]	Africa	Landsat y Sentinel	36,924	RF	2	11	Espectrales, Topograficas	94.5 %
2018	[86]	China	Landsat	2	MCC	9	9	Espectrales	93.05 %
2018	[12]	EUA	Landsat	1322	DNN	2	10	Espectrales, Temporales	97 %
2018	[23]	Africa	Landsat, Modis, WorldClim, Night-time light (NTL)	5266	RF	9	145	Espectrales, Temporales, Topográficas	76 %

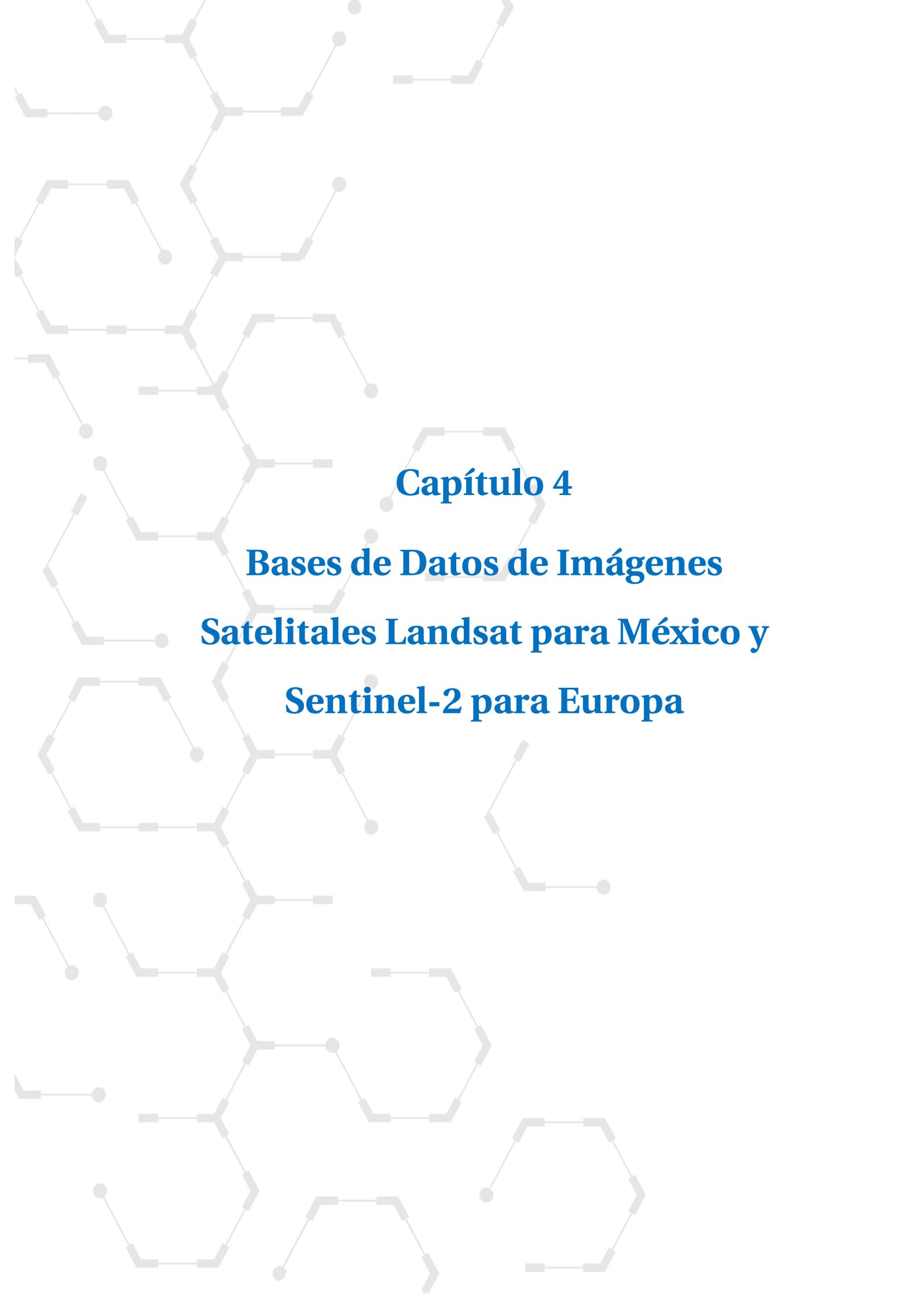
Tabla 3.1: Resumen de trabajos

Año	Referencia	Región	Sensor	Imágenes	Modelo	Clases	Características		O.A.
							Total	Tipos	
2018	[49]	Mongolia	Landsat	8	SVM	8	6	Espectrales	92%
2018	[87]	Etiopia	Landsat	12	MLC	6	14	Espectrales	91.8%
2018	[76]	Brasil	Modis	13,800	SVM	9	4	Espectrales, Temporales	94%
2018	[61]	Ucrania	Landsat, Sentinel-1A, Sentinel 2	41	ANN	13	12	Espectrales, Radar	92.4%
2017	[18]	Global	Landsat	—	OBIA	4	—	Espectrales	80%-97%
2018	[85]	EUA	Landsat	26	SVM	11	6	Espectrales	84.7%
2017	[50]	Sudáfrica	Landsat	5	SVM	7	33	Espectrales, Estadísticas, Textura	96%
2012	[93]	España	Landsat	2	RF	14	16	Espectrales, Topográficas	92%
2018	[21]	Brasil	Landsat	2200	RF	2	10	Espectrales	96%
2017	[24]	Africa	Landsat	5141	RF	6	13	Espectrales, Estadísticas	89%
2018	[75]	Europa, Africa y M. Oriente	Landsat	2167	LDA	2	10	Espectrales, Estadísticas, Topográficas	74%
2018	[84]	Malasia	Landsat	4	ANN	5	6	Espectrales	98.22%
2018	[66]	China	Landsat	11597	RF	2	19	Espectrales, Estadísticas, Topográficas, Geográficas	90%

Tabla 3.1: Resumen de trabajos (cont.)

Año	Referencia	Región	Sensor	Imágenes	Modelo	Clases	Características		O.A.
							Total	Tipos	
2016	[39]	India	AVIRIS	1	SVM	5	220	Espectrales, Estadísticas	92.6%
2017	[29]	Africa	Worldview-3	1	SVM	10	169	Espectrales, Estadísticas, Textura	80.1%
2018	[105]	China	GaoFen-2, Beijing-2, UAV	3	SVM	5	180	Espectrales, Estadísticas, Textura, Geométricas, Geográficas, Indices de Sombra	89.85%
2016	[95]	China	UAV	1	fuzzy classification	7	34	Espectrales, Estadísticas, Textura, Forma	81.6%

Tabla 3.1: Resumen de trabajos (cont.)



Capítulo 4
Bases de Datos de Imágenes
Satelitales Landsat para México y
Sentinel-2 para Europa

Capítulo 4 Bases de Datos de Imágenes Satelitales Landsat para México y Sentinel-2 para Europa

En este trabajo se utilizaron dos tipos de imágenes: Landsat y Sentinel-2. En el caso de Landsat, se llevó a cabo la construcción del conjunto de datos a partir de una colección nacional de 3,515 imágenes correspondientes al año 2011, por lo que en este capítulo se explica a detalle el proceso realizado para la integración de un mosaico nacional libre de nubes y el proceso para contar con parches de píxeles etiquetados.

En el caso de Sentinel-2, se utilizó una base de datos publicada y etiquetada por otro equipo de investigadores, por lo que indicamos la fuente y las características generales de los datos usados.

4.1. Imágenes Landsat

El conjunto de datos Landsat proviene del continuo nacional libre de nubes generado para este trabajo y construido a partir de la geomédiana calculada para el año 2011; consta de 60,000 imágenes seleccionadas al azar de 33×33 píxeles. Usando la plataforma Open Data Cube (ODC) [27, 64] indexamos las 3,515 imágenes disponibles para México del año 2011, las cuales corresponden a múltiples coberturas de todo el país. A partir de lo cual, se generó un compuesto de píxeles anual libre de nubes de todo México usando la mediana geométrica (geomédiana). Como se define en Roberts et al. [77], la geomédiana produce un píxel-resumen para cada región de $30\text{m} \times 30\text{m}$ del país a partir de las múltiples imágenes captadas a lo largo del año, lo que resulta en las primeras seis bandas en un mosaico nacional (azul, verde, rojo, NIR, SWIR1 y SWIR2), solamente se incorporan esas seis bandas en las imágenes almacenadas en el ODC.

Se calcularon más bandas para representar diferentes condiciones ambientales

durante 2011, dichas condiciones se pueden captar analizando diversas combinaciones de bandas, a estas combinaciones también se les llaman índices espectrales, los cuales permiten detectar vegetación, presencia de agua u otro tipo de coberturas. Gracias a la tecnología del ODC se tiene una serie de tiempo de píxeles alineados espacialmente. Por lo que para cada región de $30\text{m} \times 30\text{m}$ que cubre un píxel, se tienen múltiples píxeles a lo largo del tiempo, particularmente en este manuscrito se tomaron los correspondientes al año 2011, dando como resultado una serie de tiempo de píxeles con seis bandas espectrales cada uno, con una frecuencia que depende de las imágenes disponibles para México en el acervo Landsat.

Entonces, para caracterizar el cambio a lo largo del año, se generaron 15 índices espectrales que pueden calcularse para cada instancia de píxel en la serie de tiempo para la región que abarca cada uno de los píxeles que cubren al país. Es decir, si en una de estas regiones se cuenta con 20 píxeles recolectados a lo largo del año, a cada uno de esos 20 píxeles se les puede calcular 15 índices espectrales.

Dichos índices corresponden a todas las combinaciones posibles tomando dos bandas para el conjunto de las seis disponibles para Landsat. Cada combinación de dos bandas se calculó usando el patrón de diferencias normalizadas, que se muestra más adelante.

Para resumir a los píxeles recolectados de cada región de $30\text{m} \times 30\text{m}$ del país a lo largo del año, se calcularon cinco estadísticos (mínimo, promedio, máximo, desviación estándar y mediana). Lo que nos deja 15 índices espectrales resumidos cada uno con cinco estadísticos. Es decir, si tenemos el índice espectral NDVI, que es la combinación de las bandas NIR y rojo, calculado para los hipotéticos 20 píxeles mencionados anteriormente. Lo que da 20 valores NDVI para la región de $30\text{m} \times 30\text{m}$.

Para resumir, todo el año se decidió calcular los 5 estadísticos sobre la serie de valores de dicho índice espectral, es decir el mínimo NDVI, el NDVI promedio y así subsecuentemente para los 5 estadísticos, por lo que tendremos cinco valores que resumen la combinación nir-rojo de la serie de tiempo anual de los píxeles recolectados para la región correspondiente a dicho píxel. Los valores resultantes de aplicar el mis-

mo procedimiento a el resto de las combinaciones de dos bandas, se agregaron como nuevas bandas a las 6 originales provenientes del resumen geomédiana. Esto es, 15 índices espectrales resumidos cada uno con los cinco estadísticos nos dan 75 nuevas bandas. Finalmente, usando los seis valores de la geomédiana se calcularon 21 bandas más a partir de índices espectrales bien conocidos en la literatura, dando como resultado: 6 bandas de la geomédiana + 75 bandas del análisis de las series de tiempo de píxeles + 21 índices espectrales derivados de la geomédiana nos da un total de 102 bandas cubriendo todo México.

Luego de lo anterior, se seleccionaron al azar 5000 regiones de 33 píxeles \times 33 píxeles de cada clase LULC estudiada en este trabajo. Para esta selección aleatoria se tomaron las regiones de 33 px \times 33 px que tienen 70% o más de sus píxeles de una clase única. Lo anterior, para buscar tener datos consistentes y confiables para la tarea de clasificación automática. Esta selección aleatoria genera un total de 60,000 regiones, los datos están disponibles para su descarga en la siguiente referencia [16]. La Tabla 4.1 describe el conjunto de datos, por ejemplo, el número de muestras seleccionadas para cada clase y una descripción general de cada una. Aquí, se puede ver el nombre de la clase, la descripción general, el número de imágenes muestreadas y el área correspondiente en kms².

4.1.1. Procedimiento de construcción del conjunto de datos

Como punto de partida, se tomó la información de uso de suelo y vegetación producida por el INEGI y la cual está disponible en el Catálogo de metadatos geográficos de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) [42]. Los datos usados se describen de la siguiente forma en el documento de metadatos: “Los Conjuntos de Datos Vectoriales de Uso del Suelo y Vegetación, Escala 1:250 000 - Serie V, contiene información del Uso del Suelo y Vegetación obtenida a partir de la aplicación de técnicas de fotointerpretación con imágenes de satélite Landsat TM5 seleccionadas del año 2011. Esta interpretación está apoyada con trabajos de campo. Los Conjuntos de Datos contienen la ubicación, distribución y extensión de diferentes comunidades vegetales y usos agrícolas con sus respectivas variantes en tipos

de vegetación, de usos agrícolas, e información ecológica relevante. Dicha información geográfica digital contiene datos estructurados en forma vectorial codificados de acuerdo con el Diccionario de Datos Vectoriales de Uso del Suelo y Vegetación Serie IV para la Escala 1:250 000 aplicables a las diferentes unidades ecológicas (comunidades vegetales y usos antrópicos) contenidos en el conjunto de datos.” Por lo que el archivo vectorial utilizado en este manuscrito corresponde a todo México, el cual está etiquetado manualmente mediante el siguiente proceso:

Interpretación preliminar: Usando imágenes del satélite Landsat con la combinación de bandas Rojo, Verde y Azul, se realiza la identificación manual de polígonos de cambio usando como base la información de la versión anterior de uso de suelo y vegetación, con lo cual se identifican los puntos de interés a ser verificados en campo. A partir de los cuales se definen los itinerarios para el recorrido de verificación y se plantean hipótesis sobre las áreas de cambio.

Verificación de Campo: Se realizan recorridos de campo en los puntos de interés con el fin de recopilar información cualitativa y cuantitativa respecto al estado en el que se encuentra la vegetación y las características específicas del terreno.

Análisis e Integración: En esta etapa se analiza la información obtenida de las etapas anteriores y se validan o descartan las hipótesis planteadas en la interpretación preliminar a partir de los resultados obtenidos en la verificación de campo. Con lo que se actualizan los polígonos identificados originalmente.

Edición: La información actualizada se integra en un conjunto nacional de datos en formato vectorial.

Como parte del proceso, se realiza un trabajo exhaustivo en la identificación de las clases reportadas en la información de uso de suelo y vegetación. Dichas clases se organizan de acuerdo con las características de los diferentes tipos de vegetación. El INEGI identifica 58 clases de vegetación, 21 clases que agrupan diferentes tipos de agricultura, actividades acuícolas y forestales, así como cinco clases que incluyen información de elementos que no forman parte de la cobertura vegetal pero que inciden sobre ellas como las áreas urbanas, agua superficial y áreas sin vegetación. Dando un

total de 83 clases principales. Además de las clases principales, el INEGI identificó y mapeó 95 variaciones de las 58 clases de vegetación en función de la respuesta que han generado las comunidades vegetales a partir de disturbios como incendios, huracanes, inundaciones, plagas, variaciones climáticas, etcétera, dando un total de 178 clases en la Serie V de información de uso de suelo y vegetación. Lo cual muestra el detalle con el que se genera la información y explica la razón por la que llegan a pasar varios años entre las actualizaciones de dicho conjunto de datos.

Como producto, se obtiene el siguiente archivo vectorial en formato Shapefile ¹. El cual fue tomado para construir la base de datos Landsat usada en el presente documento.

Con la información descargada en coordenadas Geográficas se realizó una transformación a la proyección cónica equiáreas de Albers con los siguientes parámetros: Falso Este: 2500000.0; Falso Norte: 0.0; Longitud del origen: -102.0°; Latitud del origen: 12.0°; Primer paralelo estándar: 17.5°; Segundo paralelo estándar: 29.5°; Unidad lineal: Metro (1.0); Elipsoide de referencia: GRS80. La transformación se realiza con la finalidad de tener una proyección nacional con la que se facilite la integración de mosaicos completos de México.

Una vez proyectados los datos, se realizó la identificación de correspondencias entre el sistema de clasificación del INEGI que contiene el archivo, al del sistema de clasificación nivel 2 de Cobertura de Suelo usado en el trabajo de Gebhardt y otros [28], con la finalidad de hacer más comparables los resultados de este manuscrito con los publicados en dicho trabajo. En la Tabla 4.1, se encuentra la descripción de las clases usadas para el entrenamiento y validación de los datos Landsat correspondientes a México. En el caso de la Tabla 4.2 contiene las correspondencias entre las clases usadas en el entrenamiento y las clases publicadas por el INEGI en el producto denominado: Conjunto de Datos Vectoriales de Uso del Suelo y Vegetación, Escala 1:250 000 - Serie V. La Figura 4.1 muestra un mapa nacional de la clasificación nivel 2 usado en [28]. En la figura se puede observar la presencia de agricultura en todo el país, también que hay predominio de matorrales en la zona norte, así como los diversos tipos de bosques

¹<http://www.conabio.gob.mx/informacion/gis/maps/geo/usv250s5ugw.zip>

templados y en el sureste se puede apreciar la existencia de bosques tropicales.

Tabla 4.1: Descripción de clases seleccionadas para este trabajo.

Clase	Descripción	N muestras	Área km ²
Agricultura	Agricultura de Riego, Temporal y de Suelo Húmedo	5000	4950
Tierra estéril	Sin vegetación aparente	5000	4950
Pastizal	Pastizal Natural, inducido y cultivado así como sabana , vegetación de dunas costeras y praderas	5000	4950
Matorral	Todos los matorrales crasicauales, sarcocuale, coníferos y desérticos así como vegetación en galería, vegetación halófila y gipsófila, así como mezquite (desértico)	5000	4950
Bosque templado de coníferas	Todos los tipos de bosques de coníferas pino, abeto, cedro y enebros	5000	4950
Bosque templado caducifolio	Todos los tipos de bosques caducifolios templados de hoja ancha, principalmente roble	5000	4950
Bosque templado mixto	Todos los tipos de bosques templados mixtos, ya sea de roble-pino o bosque de pino-roble	5000	4950
Bosque tropical caducifolio	Todos los tipos de bosques tropicales caducifolios, como también matorrales subtropicales, submontanos y bosques de mezquite (tropicales)	5000	4950
Bosque tropical perennifolio	Todos los tipos de bosques tropicales de hoja perenne altos, medios y bajos, como también manglares, palmeras naturales, bosques de galería, bosques nublados y vegetación de peten	5000	4950
Área urbana	Asentamientos humanos	5000	4950
Cuerpo de agua	Cuerpos de agua y acuicultura	5000	4950
Vegetación de humedales	Vegetación de pantanos, llanuras de inundación y otra vegetación halófila	5000	4950

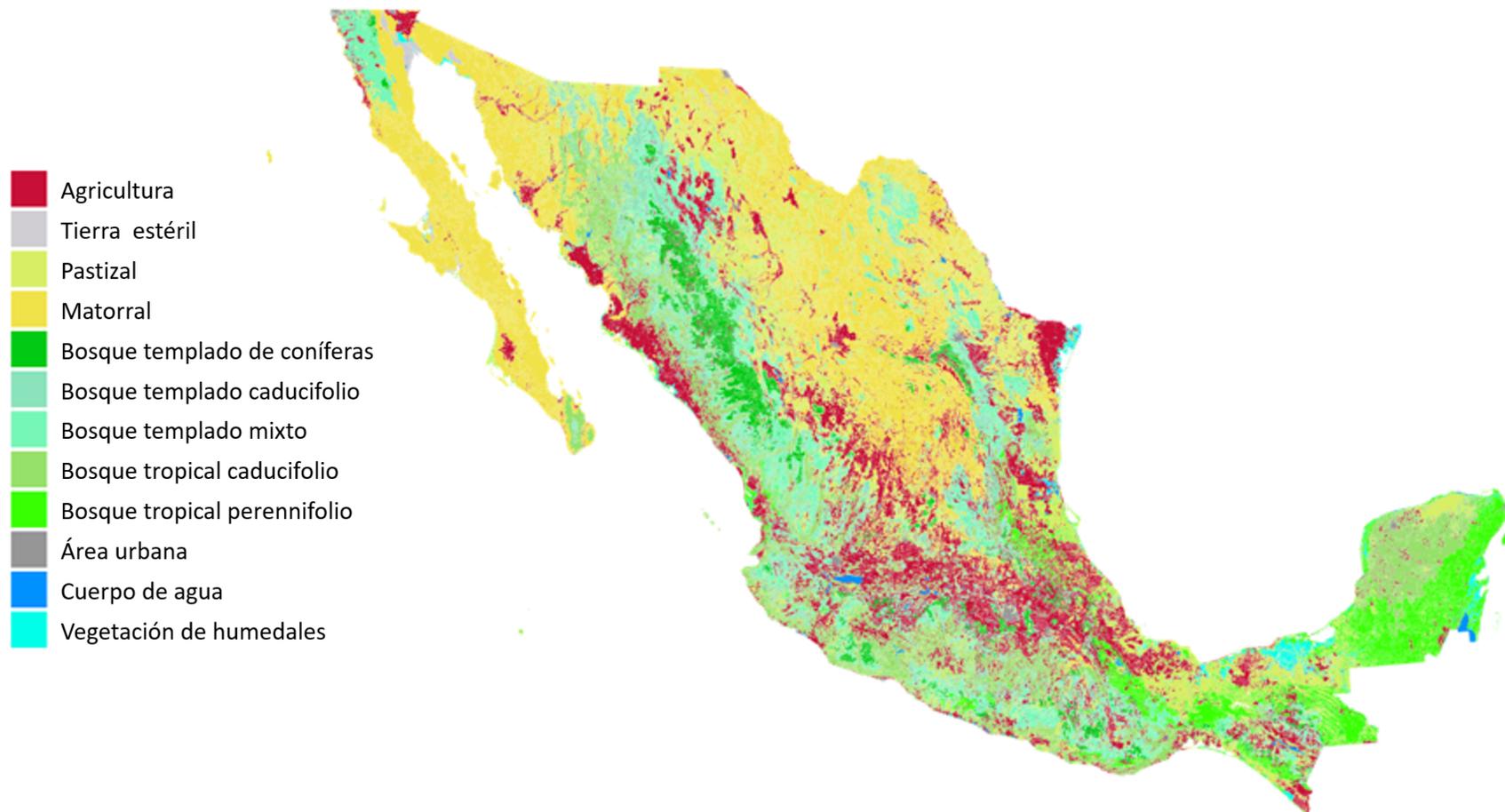


Figura 4.1: Clasificación de Cobertura de Suelo Nacional derivada a partir de la aplicación de correspondencias entre las clases de CONABIO-INEGI sobre Serie V del Uso de Suelo y Vegetación publicada por el INEGI, resultando en las 12 clases usadas en el trabajo de Gebhardt y otros [28] y también en el presente estudio, (elaboración propia).

Tabla 4.2: Correspondencias entre clases CONABIO-INEGI.

Número	Clasificación de CONABIO	Correspondencia a la clasificación INEGI
1	Bosque templado de coníferas	bosque de ayarín, bosque de cedro, bosque de oyamel, bosque de pino, bosque de táscate, vegetación secundaria arbustiva de bosque de oyamel, vegetación secundaria arbustiva de bosque de ayarín, vegetación secundaria arbustiva de bosque de cedro, vegetación secundaria arbustiva de bosque de pino, vegetación secundaria arbustiva de bosque de táscate, vegetación secundaria arbórea de bosque de ayarín, vegetación secundaria arbórea de bosque de cedro, vegetación secundaria arbórea de bosque de oyamel, vegetación secundaria arbórea de bosque de pino, vegetación secundaria arbórea de bosque de táscate, vegetación secundaria herbácea de bosque de pino, vegetación secundaria herbácea de bosque de táscate, vegetación secundaria herbácea de matorral de coníferas
2	Bosque templado caducifolio	bosque de encino, vegetación secundaria arbustiva de bosque de encino, vegetación secundaria arbórea de bosque de encino, vegetación secundaria herbácea de bosque de encino
3	Bosque templado mixto	bosque cultivado, bosque de encino-pino, bosque de galería, bosque de mezquite, bosque de pino-encino, bosque inducido, chaparral, matorral submontano, mezquital tropical, vegetación secundaria arbustiva de bosque de encino-pino, vegetación secundaria arbustiva de bosque de galería, vegetación secundaria arbustiva de bosque de mezquite, vegetación secundaria arbustiva de bosque de pino-encino, vegetación secundaria arbustiva de chaparral, vegetación secundaria arbustiva de matorral submontano, vegetación secundaria arbustiva de mezquital tropical, vegetación secundaria arbórea de bosque de encino-pino, vegetación secundaria arbórea de bosque de galería, vegetación secundaria arbórea de bosque de mezquite, vegetación secundaria herbácea de bosque de pino-encino, vegetación secundaria herbácea de bosque de pino-encino
4	Bosque tropical perennifolio	bosque mesófilo de montaña, manglar, palmar inducido, palmar natural, selva alta perennifolia, selva alta subperennifolia, selva baja espinosa subperennifolia, selva baja perennifolia, selva baja subperennifolia, selva de galería, selva mediana perennifolia, selva mediana subperennifolia, vegetación de petén, vegetación secundaria arbustiva de bosque mesófilo de montaña, vegetación secundaria arbustiva de manglar, vegetación secundaria arbustiva de palmar natural, vegetación secundaria arbustiva de selva alta perennifolia, vegetación secundaria arbustiva de selva alta subperennifolia, vegetación secundaria arbustiva de selva baja espinosa subperennifolia, vegetación secundaria arbustiva de selva baja perennifolia, vegetación secundaria arbustiva de selva mediana perennifolia, vegetación secundaria arbustiva de selva mediana subperennifolia, vegetación secundaria arbórea de bosque mesófilo de montaña, vegetación secundaria arbórea de manglar, vegetación secundaria arbórea de palmar natural, vegetación secundaria arbórea de selva alta perennifolia, vegetación secundaria arbórea de selva alta subperennifolia, vegetación secundaria arbórea de selva baja espinosa subperennifolia, vegetación secundaria arbórea de selva baja perennifolia, vegetación secundaria arbórea de selva baja subperennifolia, vegetación secundaria arbórea de selva de galería, vegetación secundaria arbórea de selva mediana subperennifolia, vegetación secundaria arbórea de vegetación de petén, vegetación secundaria herbácea de bosque mesófilo de montaña, vegetación secundaria herbácea de manglar, vegetación secundaria herbácea de selva alta perennifolia, vegetación secundaria herbácea de selva alta subperennifolia, vegetación secundaria herbácea de selva baja espinosa subperennifolia, vegetación secundaria herbácea de selva mediana subperennifolia
5	Bosque tropical caducifolio	matorral subtropical, selva baja caducifolia, selva baja espinosa caducifolia, selva baja subcaducifolia, selva mediana caducifolia, selva mediana subcaducifolia, vegetación secundaria arbustiva de matorral subtropical, vegetación secundaria arbustiva de selva baja caducifolia, vegetación secundaria arbustiva de selva baja espinosa caducifolia, vegetación secundaria arbustiva de selva baja subcaducifolia, vegetación secundaria arbustiva de selva mediana caducifolia, vegetación secundaria arbustiva de selva mediana subcaducifolia, vegetación secundaria arbórea de selva baja caducifolia, vegetación secundaria arbórea de selva baja espinosa caducifolia, vegetación secundaria arbórea de selva baja subcaducifolia, vegetación secundaria arbórea de selva mediana caducifolia, vegetación secundaria arbórea de selva mediana subcaducifolia, vegetación secundaria herbácea de selva baja caducifolia, vegetación secundaria herbácea de selva baja espinosa caducifolia, vegetación secundaria herbácea de selva mediana caducifolia, vegetación secundaria herbácea de selva mediana subcaducifolia

Número	Clasificación de CONABIO	Correspondencia a la clasificación INEGI
6	Matorral	matorral crasicaule, matorral desértico micrófilo, matorral desértico rosetófilo, matorral espinoso tamaulipeco, matorral rosetófilo costero, matorral sarco-crasicaule, matorral sarco-crasicaule de neblina, matorral sarco-caule, mezquital xerófilo, vegetación de desiertos arenosos, vegetación de galería, vegetación gipsófila, vegetación halófila xerófila, vegetación secundaria arbustiva de matorral crasicaule, vegetación secundaria arbustiva de matorral desértico micrófilo, vegetación secundaria arbustiva de matorral desértico rosetófilo, vegetación secundaria arbustiva de matorral espinoso tamaulipeco, vegetación secundaria arbustiva de matorral rosetófilo costero, vegetación secundaria arbustiva de matorral sarco-crasicaule, vegetación secundaria arbustiva de matorral sarco-caule, vegetación secundaria arbustiva de mezquital xerófilo, vegetación secundaria arbustiva de vegetación de desiertos arenosos, vegetación secundaria arbustiva de vegetación de galería, vegetación secundaria arbustiva de vegetación halófila xerófila, vegetación secundaria herbácea de matorral crasicaule, vegetación secundaria herbácea de matorral desértico micrófilo, vegetación secundaria herbácea de matorral desértico rosetófilo, vegetación secundaria herbácea de matorral espinoso tamaulipeco, vegetación secundaria herbácea de matorral rosetófilo costero, vegetación secundaria herbácea de matorral sarco-crasicaule, vegetación secundaria herbácea de matorral sarco-caule de neblina, vegetación secundaria herbácea de vegetación halófila xerófila
7	Vegetación de humedales	popal, tular, vegetación halófila hidrófila, vegetación secundaria arbustiva de vegetación halófila hidrófila
8	Agricultura	agricultura de humedad anual, agricultura de humedad anual y permanente, agricultura de humedad anual y semipermanente, agricultura de humedad permanente, agricultura de humedad semipermanente, agricultura de humedad semipermanente y permanente, agricultura de riego anual, agricultura de riego anual y permanente, agricultura de riego anual y semipermanente, agricultura de riego permanente, agricultura de riego semipermanente, agricultura de riego semipermanente y permanente, agricultura de temporal anual, agricultura de temporal anual y permanente, agricultura de temporal anual y semipermanente, agricultura de temporal permanente, agricultura de temporal semipermanente, agricultura de temporal semipermanente y permanente
9	Pastizal	pastizal cultivado, pastizal gipsófilo, pastizal halófilo, pastizal inducido, pastizal natural, pradera de alta montaña, sabana, sabanoide, vegetación de dunas costeras, vegetación secundaria arbustiva de pastizal gipsófilo, vegetación secundaria arbustiva de pastizal halófilo, vegetación secundaria arbustiva de pastizal natural, vegetación secundaria herbácea de pastizal natural
10	Cuerpo de agua	acuícola, cuerpo de agua
11	Tierra estéril	desprovisto de vegetación, sin vegetación aparente
12	Área urbana	asentamientos humanos, zona urbana

Nota: Tabla de elaboración propia con información aportada por el Ingeniero Jesús Abad Argumedo Espinoza.

Una vez generado el mapa de 12 clases a nivel nacional, se procedió a rasterizar el mapa vectorial. Dando como resultado una imagen de 110,000 × 70,000 píxeles dónde el valor de cada píxel indica un valor entre 1 a 12 correspondiente a la clase. La imagen corresponde a todo México con píxeles de una resolución a 30 metros. Posteriormente se realizó un rejilla de 1 km × 1 km cubriendo todo México. Resultaron 1,975,940 celdas. Para cada celda se tomaron 33 × 33 píxeles completos para realizar el cálculo de la clase predominante.

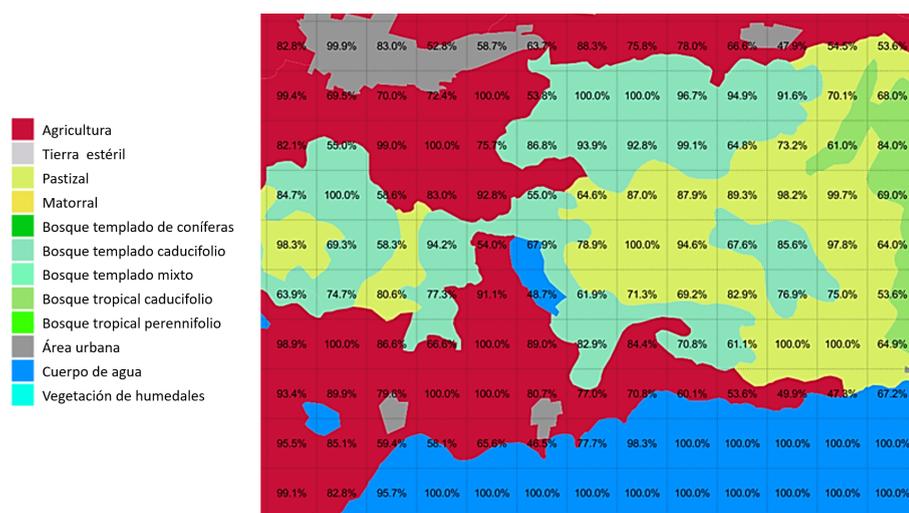


Figura 4.2: Identificación del porcentaje de píxeles de la clase mayoritaria para cada celda de 33 x 33 a partir de la Clasificación de Cobertura de Suelo Nacional resultante del mapeo de clases CONABIO-INEGI (elaboración propia).

A cada celda de 1 km × 1 km se le asoció el porcentaje correspondiente a la clase con más píxeles, en la figura 4.2 se puede ver el porcentaje asociado a la clase con mayor presencia en la región. El total de regiones de 70% o más píxeles de una clase dominante corresponde a 1,640,827 lo cual es un 83% del territorio nacional. Es decir, solamente el 17% de las regiones de 1km x 1km tienen menos de 70% de píxeles de una clase dominante. Entonces se seleccionaron aleatoriamente 5,000 regiones de cada clase a nivel nacional, considerando solamente aquellas regiones que tienen al menos un 70% de píxeles de la clase dominante. Con lo anterior se busca tener una muestra consistente para los ejercicios de clasificación, y como resultado se cuenta con 12 conjuntos, uno por cada clase, de 5,000 imágenes que dan un total de 60,000 regiones seleccionadas aleatoriamente a lo largo de todo México como se puede ver en la figura 4.3.

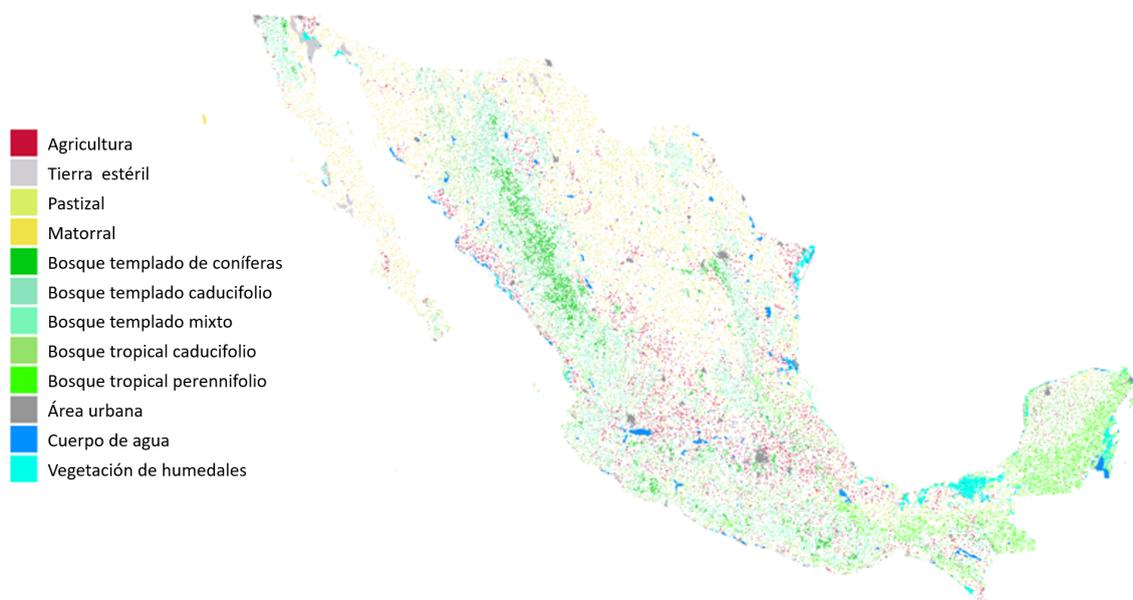


Figura 4.3: Mapa de la selección aleatoria de 5,000 celdas por clase, en cada celda predomina al menos un 70 % de píxeles de una sola clase, dando como resultado un total de 60,000 regiones, (elaboración propia).

Dado que todas las imágenes contienen píxeles de otras categorías, la tabla 4.3 muestra el porcentaje de píxeles considerados como la clase predominante y sus porcentajes de pertenencia a las otras clases. Es decir, por ejemplo, la clase 1 (Bosque templado de coníferas) se tiene un 95.2% de píxeles etiquetados como clase 1, y el 4.8% restante se distribuyó en el resto de clases siendo la más frecuente la clase 3 (Bosque templado mixto) con 1.6% de píxeles. Entonces, en esta tabla se puede analizar qué clases están más y menos mezcladas y con cuáles. Hasta aquí, se cuenta con regiones de 1 km × 1 km etiquetadas. El siguiente paso implica la incorporación de información de las imágenes de satélite Landsat.

Tabla 4.3: Porcentaje de mezcla de cada clase en el conjunto Landsat.

Clase predominante		Clase mezclada											
		1	2	3	4	5	6	7	8	9	10	11	12
1	Bosque templado de coníferas	95.2%	0.3%	1.6%	0.1%	0.1%	0.3%	0.0%	1.4%	1.0%	0.0%	0.0%	0.0%
2	Bosque templado caducifolio	0.3%	93.2%	1.9%	0.1%	1.6%	0.2%	0.0%	0.9%	1.8%	0.0%	0.0%	0.0%
3	Bosque templado mixto	0.7%	1.1%	95.0%	0.1%	0.3%	0.6%	0.0%	1.1%	1.0%	0.0%	0.0%	0.0%
4	Bosque tropical perennifolio	0.0%	0.1%	0.2%	96.5%	0.2%	0.0%	0.3%	0.6%	1.6%	0.2%	0.1%	0.1%
5	Bosque tropical caducifolio	0.0%	0.8%	0.3%	0.1%	95.2%	0.2%	0.0%	1.8%	1.2%	0.1%	0.0%	0.1%
6	Matorral	0.0%	0.0%	0.2%	0.0%	0.0%	97.9%	0.0%	0.6%	1.0%	0.0%	0.0%	0.0%
7	Vegetación de humedales	0.0%	0.0%	0.0%	2.4%	0.1%	0.3%	93.3%	0.6%	1.5%	1.6%	0.3%	0.0%
8	Agricultura	0.3%	0.3%	0.5%	0.3%	1.0%	0.8%	0.0%	94.6%	1.4%	0.2%	0.0%	0.6%
9	Pastizal	0.1%	0.7%	0.5%	0.7%	0.8%	1.7%	0.1%	1.6%	93.5%	0.1%	0.0%	0.1%
10	Cuerpo de agua	0.0%	0.0%	0.1%	0.8%	1.0%	0.8%	0.7%	1.2%	1.3%	93.8%	0.2%	0.1%
11	Tierra estéril	0.0%	0.1%	0.0%	0.4%	0.1%	2.2%	0.3%	0.2%	0.6%	0.4%	95.5%	0.0%
12	Área urbana	0.0%	0.0%	0.4%	0.2%	0.5%	0.6%	0.0%	3.7%	1.1%	0.1%	0.0%	93.2%

4.1.2. Cubo de Datos Geospaciales de Imágenes de Satélite

Para llevar a cabo los análisis correspondientes a la región de México se utilizaron todas las imágenes Landsat disponibles para México del año 2011, el cual corresponde al año de referencia del Conjunto de Datos de la Carta de Uso de Suelo Serie V, que fue la base para crear el etiquetado de las celdas mostrado en la sección anterior. La cantidad de escenas Landsat es de 3,515, las cuales requieren 3 TB de almacenamiento. En la figura 4.4 se puede ver la distribución de las imágenes en México, cada región de la imagen representa la región que cubre cada imagen según las distribuye la NASA y el número indica la cantidad de imágenes publicadas por la Agencia Espacial para esa región durante el año 2011.

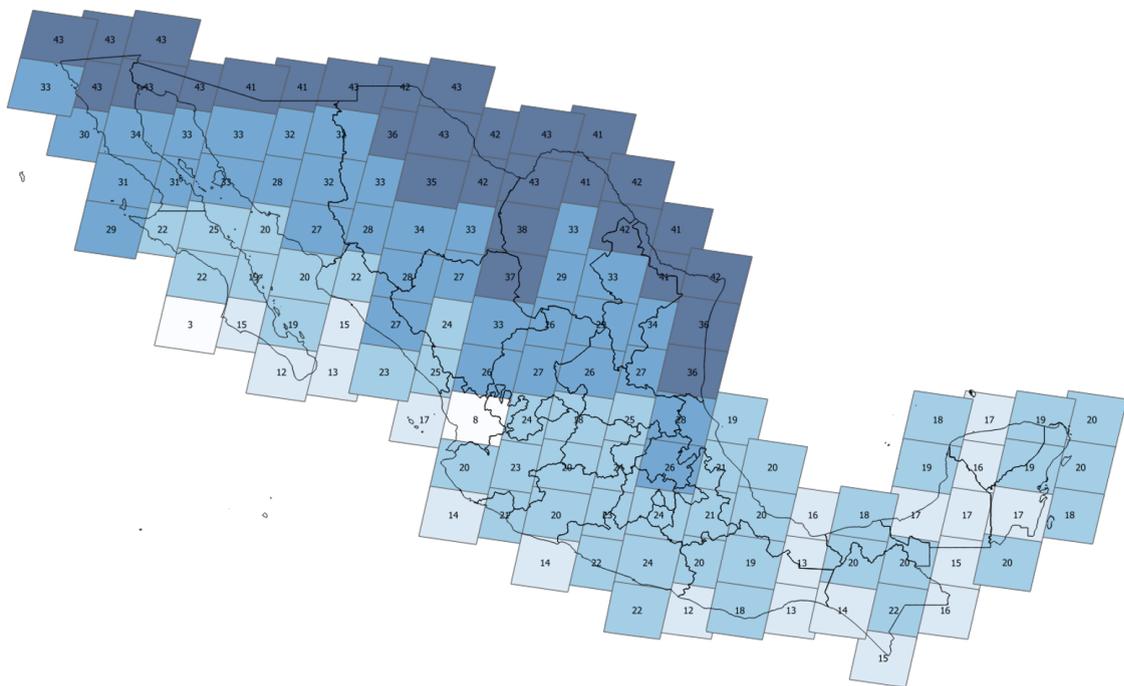


Figura 4.4: Distribución de 3,515 imágenes Landsat correspondientes al año 2011, se muestra el número de imágenes descargadas y procesadas para cada región, (elaboración propia).

Para poder manejar los grandes volúmenes de imágenes Landsat se decidió adoptar la tecnología denominada Open Data Cube (ODC) [55], desarrollado colaborativamente por parte de organismos internacionales y que busca potenciar el efecto de la información satelital al brindar un recurso de uso libre y accesible para la utilización abierta, así como impulsar una comunidad para desarrollar, mantener e incrementar

la cantidad de aplicaciones. Se llevó a cabo la instrumentación del ODC utilizando las imágenes Landsat de 2011 proporcionadas por NASA/USGS.

La implementación del ODC para esta tesis fue llevada a cabo en el INEGI, y el proceso seguido se encuentra documentado en [2]. Para el año 2011 se encontró que el 92.2% de las imágenes disponibles para la región de México en el año 2011 pertenecen a la Colección 1 de la NASA y tienen un nivel de calidad de Nivel 1 (T1), que se caracteriza por tener imágenes corregidas y coherentes con un error cuadrático medio máximo de 12 metros. El 7.8% restante corresponde al Nivel 2 (T2), que también están calibradas, pero no cumplen con los estándares de calidad del Nivel 1 debido a un error cuadrático medio mayor a 12 metros.

El ODC consiste en un conjunto de librerías de Python y un sistema de almacenamiento de datos utilizando PostgreSQL. Específicamente, el software del ODC cuenta con funciones que permiten catalogar, indexar y procesar imágenes satelitales en grandes cantidades, en el formato en que son descargadas de la agencia espacial. Es importante mencionar que la base de datos del ODC no almacena los recursos satelitales en sí, sino únicamente registra la ubicación en el almacenamiento compartido en red. Además, los metadatos de cada imagen son registrados en la base de datos, los cuales incluyen información sobre su fuente, fecha de captura, región geográfica y bandas espectrales, entre otros datos relevantes para la generación de productos derivados. De esta manera, el ODC se convierte en una herramienta que facilita el acceso y manejo de grandes cantidades de imágenes satelitales a través de programación en Python.

Los datos usados en este documento corresponden a las imágenes del satélite Landsat 5 y Landsat 7 y se tienen incorporadas las bandas Azul, Verde, Rojo, Infrarrojo Cercano, Infrarrojo de Onda Corta 1 e Infrarrojo de Onda Corta 2, todas a 30 metros de resolución. Cada píxel en el ODC tiene 6 bandas espectrales recuperadas a partir de la imagen Landsat original, de la cual proviene cada uno.

El ODC solamente integra 6 bandas debido a que está diseñado para guardar la consistencia con todas las imágenes Landsat disponibles desde los años ochentas, por

lo que se sacrifican algunas bandas disponibles en las nuevas generaciones de satélites Landsat a cambio de contar con una consistencia de varias décadas.

Las imágenes integradas en el ODC también conservan consistencia espacial, es decir, cada imagen nueva está alineada espacialmente a todas las imágenes que ya están integradas. Por lo que podemos imaginar que todos los píxeles están apilados unos sobre otros y cada pila representa una región de $30\text{m} \times 30\text{m}$.

Con las imágenes registradas en el ODC es posible acceder a ellas y procesarlas para generar gran diversidad de productos derivados del análisis de series de tiempo de imágenes de satélite. Uno de los productos que es posible generar es la geomédiana Landsat, de la cual resulta una imagen nacional con 6 bandas espectrales según se definió en el trabajo de Roberts y otros [77], que representa el estado del territorio en el periodo de análisis.

Cuando se genera la geomédiana Landsat se controlan 2 parámetros: 1) la región geográfica y 2) periodo de tiempo del análisis. De manera automática el algoritmo de la geomédiana utiliza las máscaras de calidad que acompañan a las imágenes que se incorporaron al ODC; como resultado, el algoritmo descarta los píxeles marcados en las máscaras de calidad como nube, sombra o píxeles con alta saturación.

El producto final de cada geomédiana consiste en un mosaico nacional de imágenes multispectrales exento de nubes, presentado en formato GeoTIFF. Se busca que dichos mosaicos representen las características del terreno en un periodo de tiempo específico. Para este documento se generó el mosaico nacional del año 2011 mediante el algoritmo de la geomédiana, como resultado, se generó una imagen de 35 Gb en formato TIFF del año 2011 como se puede apreciar en la figura 4.5. En la imagen se puede ver la composición final de píxeles cubriendo todo México.

A partir de la geomédiana Landsat del año 2011 se calcularon 21 índices espectrales para cada píxel en la imagen. Adicionalmente se realizaron 15 mosaicos nacionales a partir del análisis de la serie de tiempo de cada píxel disponible para el año 2011 usando todas las combinaciones normalizadas de diferencias espectrales posibles con las 6 bandas que se incorporaron al ODC. Como resultado se generaron 102 bandas las



Figura 4.5: Mosaico de la geomédiana Landsat nacional libre de nubes del año 2011, compuesta por píxeles de resolución de 30 metros x 30 metros, (elaboración propia).

cuales se encuentran listadas en la tabla 4.4

4.1.3. Publicación de la Base de Datos

Como resultado se generó una base de datos, la cual está a disposición de la comunidad en [16]. La base de datos está segmentada en 12 directorios numerados del 1 al 12, según la clase a la que pertenecen de acuerdo a la numeración de la tabla 4.2. Cada directorio contiene las imágenes de dimensiones $33 \times 33 \times 102$ en formato NPY de la librería numpy de Python. Adicionalmente se incluyó un Shapefile que contiene los polígonos de 1 km x 1 km de los cuales se extrajeron las imágenes, con los campos **MX_id** que indica el identificador único del archivo, **class** que indica la clase a la que pertenece el archivo y **percentage** que indica la proporción total de píxeles ocupados por la clase etiquetada.

Tabla 4.4: Lista de las bandas presentes en las imágenes Landsat

Orden	Num. bandas	Designación
Bandas de la geomédiana	6	azul, verde, rojo, nir, swir 1, swir 2
Índices basados en la geomédiana	15	evi, bu, sr, arvi, ui, ndbi, ibi, ndvi, ndwi, mndwi, nbi, brba, nbai, baei, bi
Transformación tasseled cap basada en la geomédiana	6	brightness, greenness, wetness, fourth, fifth, sixth
Análisis de series de tiempo por píxel del año 2011		
rojo - swir1 rojo + swir1	5	min, mean, max, std, median
rojo - nir rojo + nir	5	min, mean, max, std, median
swir1 - swir2 swir1 + swir2	5	min, mean, max, std, median
nir - swir2 nir + swir2	5	min, mean, max, std, median
nir - swir1 nir + swir1	5	min, mean, max, std, median
rojo - swir2 rojo + swir2	5	min, mean, max, std, median
verde - swir2 verde + swir2	5	min, mean, max, std, median
verde - swir1 verde + swir1	5	min, mean, max, std, median
verde - rojo verde + rojo	5	min, mean, max, std, median
verde - nir verde + nir	5	min, mean, max, std, median
azul - swir2 azul + swir2	5	min, mean, max, std, median
azul - swir1 azul + swir1	5	min, mean, max, std, median
azul - rojo azul + rojo	5	min, mean, max, std, median
azul - nir azul + nir	5	min, mean, max, std, median
azul - verde azul + verde	5	min, mean, max, std, median
Total	102	

4.2. Imágenes Sentinel-2

Para los experimentos con el sensor Sentinel-2, utilizamos los datos publicados por Helber et al. [36], que ofrece una muestra aleatoria de 27,000 parches de imágenes de 64×64 píxeles distribuidos en el territorio de Europa. Este conjunto de datos ha sido etiquetado manualmente por los autores, lo que resulta en una muestra equilibrada de 10 clases de uso y cobertura del suelo. La Tabla 4.5 muestra esta descripción del conjunto de datos, es decir, el total de imágenes que construyen este conjunto de datos, así como el total de muestras para cada clase.

Para nuestro trabajo, no hubo análisis de series temporales para las imágenes de Sentinel-2, los autores solo etiquetaron una imagen para cada región, por lo que solo usamos los parches 64×64 con sus 13 bandas. Eso significa que no fue posible generar más bandas sintéticas derivadas de series de tiempo como lo hicimos con las imágenes

nes Landsat usando el ODC. Sin embargo se generaron múltiples índices como se explica en la fase de aumento espectral del siguiente capítulo. Las imágenes Sentinel-2 están clasificadas en 10 clases: Cultivo anual, Bosque, Vegetación herbácea, Autopista, Industrial, Pastizal, Cultivo permanente, Residencial, Río, Mar y Lago. En la figura 4.6 se puede observar la distribución de las imágenes Sentinel-2, se puede observar que existe cultivos distribuidos a lo largo de toda Europa, una concentración de vegetación herbácea en el Reino Unido y hacia el oriente del mapa se concentran algunas regiones de Bosque.

Tabla 4.5: Áreas de entrenamiento y validación de Europa

Clase	Descripción	N muestras	Area km ²
Cultivo anual	Agricultura de Riego, Temporal y de Suelo Húmedo	3000	1228
Bosque	Tipos de bosques europeos	3000	1228
Vegetación herbácea	Hierba degradada	3000	1228
Autopista	Carretera pública o privada que utiliza pavimento de asfalto y concreto	2500	1024
Industrial	Edificios Industriales	2500	1024
Pastizal	Pastizales naturales, inducidos y cultivados	2000	819
Cultivo permanente	Cultivos leñosos (p. Ej., Frutales, viñedos u olivares)	2500	1024
Residencial	Edificios Residenciales	3000	1228
Río	Corriente natural de agua que fluye en un canal	2500	1024
Mar y Lago	Grandes cuerpos de agua	3000	1228

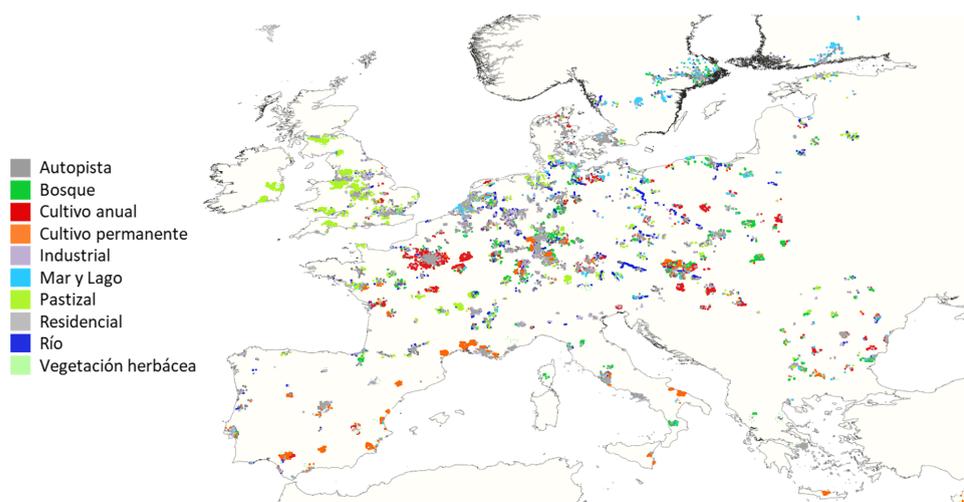


Figura 4.6: Distribución de las clases de estudio a lo largo de Europa, en una muestra de 27,000 imágenes de 64 píxeles x 64 píxeles, (elaboración propia).

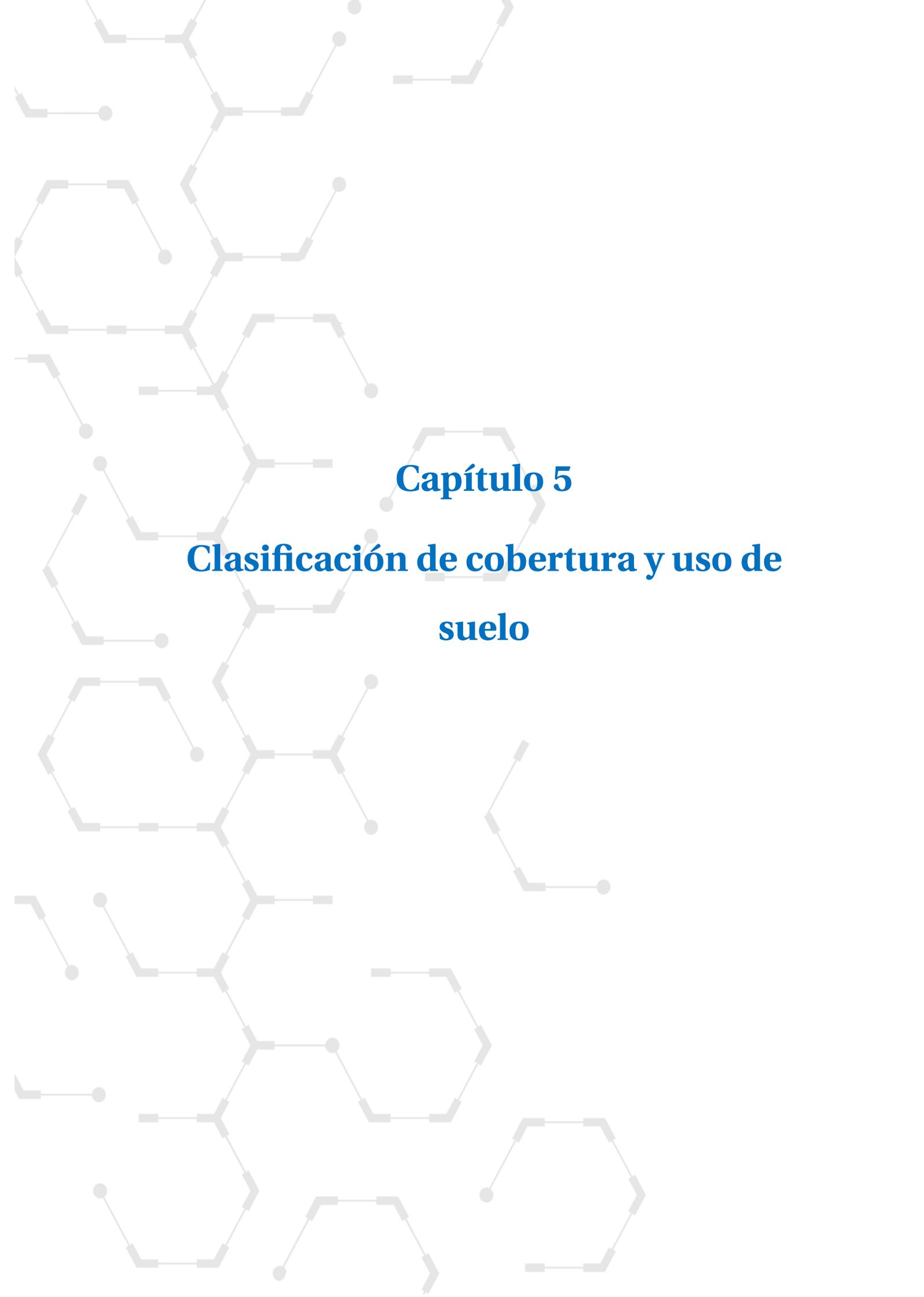
4.3. Resumen

En este capítulo se explicó el proceso de generación de la base de datos de imágenes de satélite Landsat del año 2011, la cual consta de 60,000 imágenes seleccionadas aleatoriamente de un tamaño de 33×33 píxeles con 102 bandas. Las imágenes se generaron utilizando la plataforma Open Data Cube (ODC), a partir de diferentes procesos de análisis de series de tiempo de imágenes satelitales. El etiquetado de dichas imágenes se realizó utilizando la información oficial referida al mismo año de estudio.

La información usada como base para el etiquetado se denomina: “Conjunto de datos vectoriales de uso de suelo y vegetación serie V”, publicada como dato oficial por el INEGI. Originalmente dicho conjunto de datos consiste en 178 clases a partir de las cuales se aplicaron una serie de correspondencias para disminuir el número de clases a 12, según el documento publicado por la CONABIO [28]. Lo anterior con la finalidad de realizar una comparación de los resultados logrados por la metodología propuesta en este trabajo.

La selección aleatoria de las imágenes etiquetadas se realizó tomando en cuenta que la proporción de la clase predominante en regiones de 1 km por 1 km (aproximadamente 33×33 píxeles) a lo largo de todo el país fuera de 70%. Con lo cual se busca tener una muestra consistente con las clases predominantes bien definidas. Se seleccionaron 5,000 imágenes para cada una de las 12 clases de estudio.

La base de datos resultante se publicó y quedó a disposición de la comunidad en la plataforma Zenodo [16]. En el caso de las imágenes Sentinel-2 se utilizaron datos publicados en [36], que consiste en una muestra de 27,000 parches de imágenes de 64×64 píxeles y 13 bandas espectrales distribuidos por toda Europa. Las imágenes se encuentran etiquetadas en 10 clases de cobertura de suelo con un promedio de 2,700 imágenes por clase.



Capítulo 5
Clasificación de cobertura y uso de
suelo

Capítulo 5 Clasificación de cobertura y uso de suelo

Este trabajo se centra en dos tipos de imágenes espectrales: Landsat y Sentinel-2, y dos regiones, México y Europa. Para imágenes de Landsat, utilizamos un conjunto de bandas sintéticas creadas a partir del análisis de nivel de píxeles utilizando Open Data Cube (ODC) [27, 64]. El ODC es una plataforma para el manejo y análisis de series de tiempo de imágenes satelitales. Como estrategia general para ambos conjuntos de imágenes multiespectrales, se aumenta el número de bandas en la imagen original, calculando índices espectrales muy utilizados en la comunidad de Percepción Remota. A partir de la imagen aumentada, se calculan resúmenes estadísticos, de textura y de forma. Lo que resulta en múltiples grupos de características entre los cuales se lleva a cabo una comparación de relevancia mediante la evaluación del desempeño en el proceso de clasificación LULC. Esta evaluación se realiza comparando la exactitud lograda por cada grupo de características usando múltiples métodos de clasificación supervisada.

La metodología propuesta se basa en la extracción de características para la caracterización de imágenes multiespectrales. Primeramente, los canales en las imágenes se incrementan calculando algunos índices espectrales tradicionales por píxel. De cada canal, también se calculan resúmenes estadísticos de la información y varios tipos de características relacionadas con la textura y la forma. Por lo tanto, a partir de cada parche de imagen, se calcularon automáticamente miles de características, usando la información de las bandas sintéticas extraídas a partir del ODC para Landsat y la información de las bandas originales para Sentinel-2. En este trabajo, todos los cálculos se realizaron en una estación de trabajo Dell Precision 7820, con un procesador Intel Xeon Gold 6230 con 2.1 GHz y 256 GB de RAM.

En las siguientes secciones, se explica a detalle el cálculo de cada conjunto de

características, y también el esquema de clasificación.

5.1. Extracción de características

El proceso de extracción de características se lleva a cabo para cada parche de imagen, que tiene las dimensiones de 33 x 33 píxeles para Landsat y 64 x 64 píxeles para Sentinel, el cual comienza con un aumento multiespectral. Este proceso incrementa el número de bandas originales de las imágenes usando índices espectrales para cada píxel; en el caso de Sentinel-2 se utilizaron las bandas originales de cada imagen para el cálculo de múltiples índices y en el caso de Landsat se usó la geomédiana anual para dicho proceso. Adicionalmente, gracias a que se cuenta con todas las imágenes Landsat del año 2011 para México, se lleva a cabo el análisis de la serie temporal anual de las combinaciones diferenciales de las bandas como se explicó en el capítulo anterior.

Después del proceso de aumento de bandas, se obtienen imágenes con 102 bandas para Landsat y 105 para Sentinel-2. Se aplicaron cinco estrategias de caracterización tomando cada banda individualmente, las cuales se mencionan a continuación: (1) Se generaron 17 características a partir de resúmenes estadísticos, (2) 306 características derivadas del análisis de texturas a partir de la matriz de co-ocurrencia del nivel de gris (GLCM, por sus siglas en inglés de *Gray-Level Co-occurrence Matrix*), (3) 316 que se derivan del análisis de texturas con patrones locales binarios (LBP por sus siglas en inglés de *Local Binary Patterns*) combinados con el análisis GLCM, (4) 1120 características a partir del análisis de texturas con 40 filtros de Gabor y (5) 7 características derivadas del análisis de la forma a partir de los momentos Hu para cada banda de la imagen. En otras palabras, en total se calcularon 1,766 características para cada banda de la imagen, tanto para Landsat como para Sentinel-2. Es decir, 1,766 características para cada banda (102 bandas) generan 180,132 características en el caso Landsat. Para Sentinel-2, 1,766 para cada una de las 105 bandas generan 185,430 características. Estas características están organizadas en una matriz numérica con dimensiones $m \times n$, donde m corresponde al número de parches analizados, y n , al total de características generadas por cada parche multibanda; es decir, n equivale a 180,132 para el caso Landsat y 185,430 para Sentinel-2. La figura 5.1 representa un esquema general del

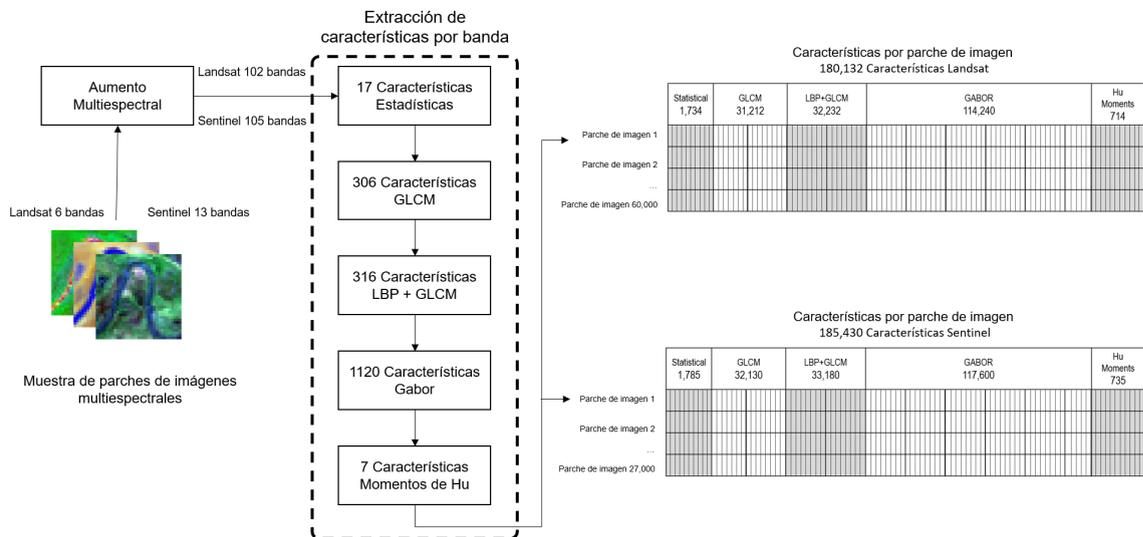


Figura 5.1: Proceso de transformación de un parche de imagen multi-banda a un renglón de características, (elaboración propia).

proceso de transformación mediante el cual se analiza cada imagen para generar la matriz de características. Por ejemplo, si tenemos una imagen Landsat como entrada, este parche tiene originalmente 6 bandas, el proceso de aumento multispectral genera un total de 102 bandas con el cálculo de varios índices espectrales y el análisis de las series de tiempo de imágenes disponibles para Landsat. Posteriormente se generan características por banda, 17 estadísticas, 306 GLCM y así hasta las 7 para los Momentos de Hu. Dando un total de 180,132 características para la imagen Landsat de entrada.

La figura 5.1 muestra la estrategia general de extracción de características para cada banda descrita anteriormente.

5.2. Aumento espectral

El proceso de aumento multispectral para Landsat consistió en tomar todas las imágenes de Landsat de 2011 y realizar análisis a nivel de píxeles para construir 6 bandas sintéticas utilizando el algoritmo geomédiana de Roberts et al. [77]. A partir de esas 6 bandas, se calcularon 21 bandas más a partir de índices espectrales comúnmente utilizados en la literatura. Finalmente, se calcularon 75 bandas a partir del análisis de la serie temporal de las 15 combinaciones de diferencias normalizadas, lo cual se

explicó a detalle en el capítulo anterior, es decir, se construyeron un total de 102 bandas. En la Tabla 4.4 del capítulo anterior se muestra la lista de cada uno de los índices usados en los datos Landsat. Este conjunto de datos generado está disponible para su descarga en Coronado, Moctezuma [16].

Para Sentinel-2 tomamos las 13 bandas originales, luego calculamos 11 índices: ARVI, BAEI, BI, BRBA, BU, EVI, IBI, NBAI, NBI, SR y UI como se muestra en la Tabla 5.1. Las características de brillo, verdor, humedad de la transformación Tasseled-Cap [71] también se calcularon, dando 3 bandas más.

Tabla 5.1: Índices espectrales utilizados en la generación de características

Índice	Referencia	Fórmula	Descripción
Atmospherically resistant vegetation index (ARVI)	[53]	$(NIR - (2 \times RED) + BLUE)/(NIR + (2 \times RED) + BLUE)$	Índice para el monitoreo de la vegetación con mejoras que disminuyen la interferencia de la atmósfera.
Built-up area extraction index (BAEI)	[9]	$(RED + 0.3)/(GREEN + SWIR1)$	Índice que combina las bandas Roja, Verde y SWIR1 para la detección de áreas urbanas.
Bare soil index (BI)	[96]	$[(SWIR1 + RED) - (NIR + BLUE)]/[(SWIR1 + RED) + (NIR + BLUE)]$	Índice para la detección de suelo descubierto.
Band ratio for built-up area (BRBA)	[97]	$RED/SWIR1$	Índice que combina las bandas Roja y SWIR1 para la detección de áreas urbanas.
Built-up Index (BU)	[104]	$[(SWIR1 - NIR)/(SWIR1 + NIR)] - [(NIR - RED)/(NIR + RED)]$	Índice que combina las bandas Roja, Infrarroja y SWIR1 para la detección de áreas urbanas.
Enhanced Vegetation Index (EVI)	[40]	$2.5 \times (NIR - RED)/(NIR + 6 \times RED - 7.5 \times BLUE + 1)$	Índice de vegetación que considera múltiples condiciones atmosféricas y la cobertura de suelo bajo la vegetación.
Index-based built-up index (IBI)	[102]	$(SWIR2 + 2 \times NIR + RED - GREEN)/(SWIR2 + 2 \times NIR + RED + GREEN)$	Índice que combina las bandas Roja, Verde, Infrarroja y SWIR2 para la detección de áreas urbanas.
Normalized built-up area index (NBAI)	[97]	$(SWIR2 - SWIR1/GREEN)/(SWIR2 + SWIR1/GREEN)$	Índice que combina las bandas Verde, SWIR1 y SWIR2 para la detección de áreas urbanas y suelo descubierto.
New built-up index (NBI)	[51]	$RED \times SWIR1/NIR$	Índice que combina las bandas Roja, Infrarroja y SWIR1 para la detección de áreas urbanas.
Simple ratio (SR)	[52]	NIR/RED	Índice que combina las bandas infrarroja y roja en una razón, útil para la detección de vegetación.
Urban Index (UI)	[54]	$(SWIR2 - NIR)/(SWIR2 + NIR)$	Índice que combina las bandas Infrarroja y SWIR2 para la detección de áreas urbanas.

Continuando con Sentinel-2, evaluamos todas las combinaciones posibles tomando 2 bandas aplicando el patrón de diferencias normalizadas. Los índices basados en diferencias normalizadas (NDI) corresponden a la mezcla de 2 bandas espectrales del mismo píxel (Banda1 y Banda2), calculando la diferencia de los valores de las bandas involucradas dividido por la suma de éstas, como se muestra en el fórmula 5.1. Tradicionalmente se han utilizado en la percepción remota para analizar y clasificar la cobertura del suelo [19].

$$NDI = \frac{\text{Banda1} - \text{Banda2}}{\text{Banda1} + \text{Banda2}} \quad (5.1)$$

Entonces, utilizando las 13 bandas espectrales originales que registra el satélite, se calculan todas las combinaciones NDI dando como resultado 78 índices adicionales, lo que resultó en 13 bandas naturales + 14 índices + 78 combinaciones NDI resultando en 105 bandas para Sentinel-2.

En resumen, se generó una imagen aumentada. Para Landsat la imagen quedó de 102 bandas y para Sentinel-2 de 105 bandas.

5.3. Características basadas en resúmenes estadísticos

La primera estrategia de generación de características consiste en tomar todos los valores de cada banda correspondientes a una imagen y calcular resúmenes estadísticos con ellos. Estos valores fueron resumidos por siete descriptores estadísticos básicos como se muestra en la tabla 5.2. Se usaron 7 descriptores los cuales son: media, varianza, mediana, variación, sesgo, curtosis y entropía discreta.

Además, las frecuencias de los valores de cada banda se calcularon con un histograma de 10 barras. Por lo que cada una de esas frecuencias se consideró como un descriptor más. Por lo tanto, todo este proceso produjo 17 características estadísticas para cada banda de la imagen.

Cada banda se trata como una imagen en escala de grises, a la que se aplican los procesos de extracción de características. Por ejemplo, en el caso de la media, se toman todos los valores de la imagen en escala de grises y se calcula la media, por lo que generamos una representación simplificada de esa banda. Aplicando la media a las 102 bandas de Landsat tenemos 102 números correspondientes al cálculo de la media. Entonces, aplicar las 17 estadísticas a las bandas de Landsat nos da 17×102 igual a 1734 características. Por lo que, a partir de los parches de imagen de las 102 bandas (Landsat) y 105 (Sentinel-2), se generan las 17 estadísticas que resultan en 1734 valores para Landsat y 1785 para Sentinel-2. Estas características se agrupan en un conjunto

Tabla 5.2: Medidas estadísticas aplicadas a vectores numéricos en esta tesis

Medida	Fórmula	Descripción
Media	$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$	Medida de tendencia central, es la suma de todos los valores dividida por el número de observaciones en el conjunto de datos.
Varianza	$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$	Medida de dispersión con respecto al valor promedio del conjunto de datos, también conocido como el cuadrado de la desviación estándar (s).
Mediana	\tilde{x}	Medida de tendencia central, resistente a valores atípicos. Dado un conjunto de datos, se organizan las observaciones en orden de menor a mayor. Si n es impar, la mediana es el valor medio. Si n es par, la mediana es la media de los dos valores medios.
Variación	$CV = \frac{s}{\bar{x}}$	Medida de variabilidad relativa independientemente de la unidad de medida.
Sesgo	$g_1 = \frac{m_3}{m_2^{3/2}}$; donde $m_i = \frac{1}{N} \sum_{n=1}^N (x_n - \bar{x})^i$; y el ' <i>i</i> ésimo' momento central	Medida de la asimetría de los valores del conjunto de datos alrededor de la media.
Curtosis	$g_2 = \frac{m_4}{m_2^2}$	Medida del peso de las colas con respecto al resto de la distribución.
Entropía	$H_e(\mathbf{p}_x) = -\sum_x p_x \log_e p_x$	Suponga que X es una variable aleatoria que toma n valores diferentes. Sea \mathbf{p}_x la probabilidad de la distribución para X y $Prob[X = x] = p_x$

llamado *Características basadas en resúmenes estadísticos*.

5.4. Características basadas en textura

El análisis de textura busca cuantificar las propiedades en las regiones de los píxeles, intuitivamente busca caracterizar la suavidad, la aspereza, la regularidad, etc. [34]. Las características de textura son útiles para obtener información estructural de las superficies y su relación con el contexto [35]. Esta información puede ser útil para el proceso de discriminación en el enfoque de clasificación supervisada.

5.4.1. Matriz de co-ocurrencia del nivel de gris (GLCM, por sus siglas en inglés de *Gray-Level Co-occurrence Matrix*)

Como una forma de extraer la textura de las imágenes, se utilizó el método conocido como GLCM [35]. La matriz GLCM caracteriza la textura de la imagen calculando la frecuencia de nivel de gris de pares de píxeles a diferentes distancias y ángulos. La GLCM es una matriz bidimensional donde el número de filas y columnas corresponde a los valores posibles del nivel de gris. Cada elemento de la matriz es una probabilidad definida por $P(i, j|d, \theta)$, que expresa la probabilidad de cada par de píxeles en la dirección θ , la distancia d y el nivel de gris de la combinación i, j , ésto en una vecindad definida. Una vez que se calcula la matriz GLCM, es posible caracterizar la textura mediante la extracción de múltiples características a partir de ella, como el segundo momento angular (ASM, por sus siglas en inglés), el contraste (CON), la correlación (COR), la disimilitud (DIS), la energía (ENE) y homogeneidad (HOM). Estas características estadísticas de segundo orden se calculan como muestra la Tabla 5.3.

Tabla 5.3: Medidas de textura estadísticas calculadas a partir del GLCM

Medida	Fórmula	Descripción
Segundo Momento Angular (ASM)	$\sum_{i,j=0}^{levels-1} P_{i,j}^2$	Mide la uniformidad de textura.
Contraste (CON)	$\sum_{i,j=0}^{levels-1} P_{i,j}(i-j)^2$	Medida de contraste de intensidad entre un píxel y su vecino en toda la imagen.
Correlación (COR)	$\sum_{i,j=0}^{levels-1} P_{i,j} \left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$	Medida de la correlación entre un píxel y su vecino en toda la imagen.
Disimilitud (DIS)	$\sum_{i,j=0}^{levels-1} P_{i,j} i-j $	Medida del total de variación presente en la imagen.
Energía (ENE)	\sqrt{ASM}	Medida de orden de una imagen.
Homogeneidad (HOM)	$\sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1+(i-j)^2}$	Mide la proximidad espacial de la distribución de elementos en el GLCM a la diagonal.

Dónde *levels* se refiere a los niveles de gris usados en la imagen; i, j corresponden a los índices (renglones, columnas) de la matriz GLCM; $P_{i,j}$ corresponde a la probabilidad estimada de cada co-ocurrencia de par de píxeles en las celdas de la matriz GLCM con respecto a todas las combinaciones observadas en la imagen; μ_i y σ_i^2 son la media y la varianza calculadas a lo largo de los renglones de la GLCM y μ_j y σ_j^2 la media y varianza a lo largo de las columnas.

Para explorar diferentes patrones en las imágenes, utilizamos varias combinaciones de parámetros para generar varias GLCM. Estos parámetros son: **niveles gris = {64}**, en exploraciones previas se identificó que cuantificar los valores de intensidad de cada píxel a 64 niveles ofrece buenos resultados en el proceso de clasificación LULC;

ángulos = {0,45,90,135} lo cual es una estrategia común que sirve para caracterizar la región alrededor de cada píxel; **distancias** = {1,2,3}, con lo que se intenta caracterizar la textura de la vecindad a una distancia de hasta 3 píxeles; **ventana** = {33} para Landsat y **ventana** = {64} para Sentinel-2, debido a los tamaños de las imágenes con los que se cuenta en cada conjunto de datos.

Con base a la experimentación previa, decidimos usar estos parámetros para tener un número limitado de combinaciones, pero aun así contar una representación amplia de la textura de la imagen en escala de grises.

A partir de cada GLCM resultante, se calculan 6 características estadísticas de segundo orden. Teniendo en cuenta que tenemos 12 GLCM, como resultado de las combinaciones de parámetros mencionadas, obtenemos 72 características de cada banda de la imagen. En total, se generaron 7344 características para Landsat (102 bandas) y 7560 para imágenes de Sentinel-2 (105 bandas).

Como se mencionó anteriormente, para todas las bandas es posible calcular 6 estadísticas de segundo orden para cada uno de los 12 GLCM derivados de combinaciones de parámetros. Es decir, tenemos un segundo momento angular (ASM) para cada una de las 12 GLCM, y así para las 5 estadísticas de segundo orden restantes. Además, con los 12 valores de cada estadística de segundo orden, calculamos las 7 estadísticas de la sección anterior es decir, la media de los ASM, la varianza de los ASM y así sucesivamente para todas las estadísticas. Es decir cada uno de los 6 grupos de estadísticas de segundo orden se resumen mediante 7 estadísticas básicas dando 42 características para cada banda de escala de grises. Aplicando lo anterior para cada banda Landsat tenemos $42 \times 102 = 4,284$ características y para las bandas Sentinel-2 tenemos $42 \times 105 = 4,410$ características.

Continuando con la caracterización de las imágenes se construye un histograma de 12 barras usando los 12 valores provenientes de cada estadístico de segundo orden derivado de las variaciones de la GLCM. Ésto da como resultado 12 valores de frecuencia y 13 límites de las barras. Usando los valores de la frecuencia calculada en el histograma, calculamos las 7 estadísticas de la sección anterior. Lo que nos da $12 +$

$13 + 7 = 32$ características para cada estadística de segundo orden. Al aplicarlo a las 6 estadísticas de segundo orden, tenemos $32 \times 6 = 192$ características para cada banda. Resultando en 19,584 características para Landsat y 20,160 para Sentinel-2.

La concatenación de todas las características anteriores da como resultado un conjunto que se llamó *Características de GLCM*, dando un total de 31,212 características para Landsat y 32,130 características para Sentinel-2.

5.4.2. Patrones binarios locales (LBP, por sus siglas en inglés de *Local Binary Patterns*) + GLCM

Continuando con las características de textura, también utilizamos el método LBP, el cual es un operador de imagen uni-banda que transforma una imagen en otra imagen de enteros que describe la textura contenida en ella. Se usa LBP en este trabajo tal como se describe en el artículo de Ojala et al. [72]. Para capturar la textura que rodea cada píxel, se usa un radio de 3 píxeles con 24 vecinos, para un píxel ubicado en el centro (x, y) , el LBP es la secuencia binaria de la comparación de los 24 vecinos con el píxel central en un radio de 3 píxeles. Es decir, el valor LBP para cada vecino puede ser 0 o 1 si es menor o mayor que el valor del píxel central. La caracterización numérica de la secuencia resultante de 24 bits representa el valor LBP, esto se puede expresar de la siguiente manera:

$$LBP(x, y) = \sum_{n=0}^{23} 2^n s(i_n - i_{x,y}), \quad (5.2)$$

dónde:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}. \quad (5.3)$$

Donde i_n corresponde al valor de escala de grises de cada píxel vecino que rodea el píxel central e $i_{x,y}$ es el valor de escala de grises del píxel central. Resultando en una cadena binaria, determinada por la función $s(x)$. Luego, se aplicó LBP a cada banda de la imagen, con un radio de 3 que proporciona una vecindad de 24 píxeles; este proceso de filtrado produce una nueva representación de imagen para cada banda original.

Es decir, 102 y 105 para las bandas transformadas de Landsat y Sentinel-2, respectivamente.

Para cada una de esas imágenes LBP resultantes, se genera un histograma de 10 bins, tomando las frecuencias correspondientes de esos histogramas, se crean 10 características para cada banda, generando 1020 características para Landsat y 1050 para Sentinel-2. Las características restantes se calcularon siguiendo el mismo esquema de cálculo que las *Características de GLCM*, aplicado a las imágenes filtradas con LBP. Como salida, se generaron 31,212 características para Landsat y 32,130 para Sentinel-2. En resumen, todas estas características calculadas se denominan *Características LBP + GLCM*. Este conjunto agrega 32,232 características para Landsat y 33,180 para Sentinel-2.

5.4.3. Características usando filtros de Gabor

Para explotar la información de textura en la imagen, también utilizamos los filtros Gabor, originalmente propuestos por Gabor et al. [25]. Utilizamos este operador usando cinco frecuencias y ocho orientaciones, lo que resulta en un banco de cuarenta filtros Gabor como se muestra en la Figura 5.2.

En esta figura se pueden ver todas las orientaciones (ver eje x) y frecuencias (ver eje y) para cada filtro. De 102 bandas de Landsat y 105 de Sentinel-2, se aplican 40 filtros Gabor para generar 40 versiones filtradas de cada banda en la imagen. Eso significa que, para cada banda, se generan 40 variaciones derivadas de los filtros de Gabor. Como resultado, se crearon imágenes con 4,080 bandas para Landsat y 4,200 para Sentinel-2, a partir de los cuales se calcularon los siete descriptores estadísticos básicos y se generaron 28,560 características para Landsat y 29,400 para Sentinel-2.

Además, para cada imagen filtrada se calculó un histograma de 10 barras a partir del cual se utilizaron las 10 frecuencias y los límites de los 11 rangos como características, para Landsat $4,080 \times 21$ resulta en 85,680 características y para Sentinel-2 tenemos $4,200 \times 21$ da 88,200 características.

Finalmente, obtenemos un total de 114,240 para Landsat y 117,600 para Sentinel-

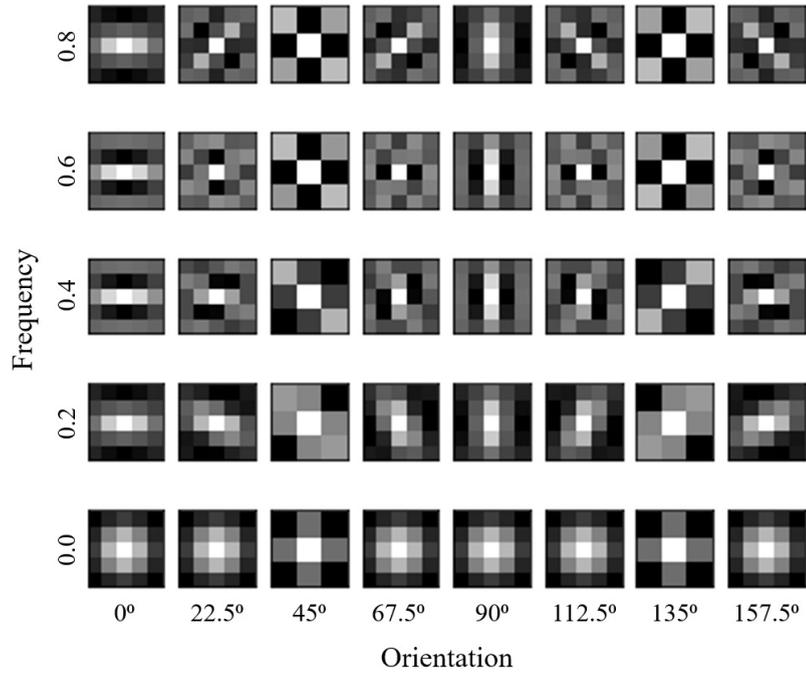


Figura 5.2: Representación visual de los 40 filtros de Gabor utilizados en el presente trabajo, (elaboración propia).

2, a este conjunto se le llamó *Características de filtros de Gabor*.

5.5. Características de forma

El análisis de forma busca caracterizar los atributos más sobresalientes que pueden identificarse en los bordes y contornos de las imágenes. Para el análisis de la forma, utilizamos los Momentos de Hu.

5.5.1. Momentos de Hu

El quinto grupo de características se llamó *Características Momentos de Hu* como se define en Hu et al. [68]. Utilizamos los momentos centrales de cada imagen, estos momentos son invariables para las transformaciones de la imagen, como la traslación, la escala y la rotación. Los Momentos de Hu se calculan de la siguiente manera:

$$H_1 = \eta_{20} + \eta_{02} \quad (5.4)$$

$$H_2 = (\eta_{20} - \eta_{02})^2 + 4(\eta_{11})^2 \quad (5.5)$$

$$H_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \quad (5.6)$$

$$H_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \quad (5.7)$$

$$H_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})((\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2) + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2) \quad (5.8)$$

$$H_6 = (\eta_{20} - \eta_{02})((\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2) + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \quad (5.9)$$

$$H_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})((\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2) + (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})(3(\eta_{30} + \eta_{12})^2 - (\eta_{03} + \eta_{21})^2). \quad (5.10)$$

Dónde:

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{00}^{(1+\gamma)}}, \gamma = \frac{p+q}{2} \quad (5.11)$$

$$\mu_{p,q} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q I(x, y) \quad (5.12)$$

$$\bar{x} = \frac{M_{10}}{M_{00}}, \bar{y} = \frac{M_{01}}{M_{00}} \quad (5.13)$$

$$M_{ij} = \sum_x \sum_y x^i y^j I(x, y) \quad (5.14)$$

Las ecuaciones desde 5.4 hasta 5.10 corresponden a los 7 momentos invariantes de Hu. La formula 5.11 corresponde a la definición del momento central normalizado el cual permite que los momentos de Hu sean invariantes a la escala el cual a su vez se basa en el momento central definidos por la fórmula 5.12. Los cuales contribuyen a que los momentos de Hu sean invariantes a la translación éstos dependen del centroide definido en 5.13 que depende de los momentos de la imagen, $I(x, y)$, como se muestra en 5.14.

Los siete momentos Hu para cada banda se calcularon generando un total de 714 características para Landsat y 735 para Sentinel-2.

En resumen, aumentamos las bandas de las imágenes calculando algunos índices espectrales que combinan múltiples bandas. A partir de las bandas resultantes se aplican diversos métodos de análisis con la finalidad de caracterizar cada banda individualmente. Se generaron características estadísticas, también se utilizó dicha estrategia estadística como un medio para caracterizar otros grupos de características. Se calcularon múltiples análisis de textura que pretenden capturar la rugosidad, la regularidad, la suavidad, etc. Y finalmente, caracterizamos la forma de la imagen a través

de momentos invariantes. Todas las características generadas se resumen en la Tabla 5.4.

Tabla 5.4: Total de características por grupo y sensor

Grupo de características	Landsat	Sentinel
Estadísticas	1,734	1,785
GLCM	31,212	32,130
LBP + GLCM	32,232	33,180
Gabor	114,240	117,600
Momentos de Hu	714	735
All features	180,132	185,430

5.6. Selección de Características y Proceso de Clasificación

Después de que se realizó el procedimiento de cálculo de características, se genera un vector muy grande (180,132 para Landsat y 185,430 para Sentinel-2) para describir cada imagen. Buscando clasificar las mejores características útiles e informativas para el problema de clasificación de uso y cobertura de suelo, se aplicó un método de selección de características.

Este método se denomina Very Large Scale ReliefF [22] (VLSReliefF), que es una variante del conocido método ReliefF [57, 92] adaptado para manejar una gran cantidad de características. Este método se utilizó porque ha demostrado un buen rendimiento con vectores de muy alta dimensión, se basa en ReliefF, que es un método de clasificación de características muy popular y muy utilizado [17, 95, 103]. Este método de selección se aplicó globalmente al conjunto de entrenamiento (70% del conjunto de datos completo) para las imágenes de ambos satélites. Como resultado, se analizaron y probaron varios porcentajes de características. Utilizamos el 10% más significativo de acuerdo al algoritmo VLSReliefF. Con dicho conjunto se llevó a cabo un análisis de la contribución de cada grupo de características.

Además, comparamos el rendimiento de cada grupo de características con cuatro algoritmos de clasificación conocidos: Random Forest [10] (RF), Support Vector Ma-

chines [70] (SVM), Extra Trees [31] (ET) y Artificial Neural Networks [37] (ANN). Se utilizó la implementación de estos algoritmos disponibles en la biblioteca scikit-learn [74].

Todos los resultados y el análisis correspondiente del ranking de características se presentan en la siguiente sección.

5.7. Resumen

En este capítulo mostramos el proceso de transformación de las bases de datos de imágenes multispectrales etiquetadas a tablas de características. La metodología empieza tomando cada imagen de la base de datos y realiza un aumento multispectral, es decir, extiende la cantidad de bandas originales a una cantidad mayor a partir del análisis y combinaciones de los datos contenidos en la imagen. El resultado de ese aumento genera 102 bandas para las imágenes Landsat y 105 bandas para las imágenes Sentinel-2.

Posterior a eso se realiza un proceso de caracterización basado en la aplicación de múltiples técnicas de extracción de características sobre cada una de las bandas de entrada en lo individual. Se empieza con la extracción de características estadísticas, las cuales son el resultado del cálculo de medidas estadísticas, como la media, la varianza, la desviación estándar y el generación de histogramas entre otras.

El siguiente tipo de características se deriva de los análisis de textura con lo que se busca caracterizar atributos en la imagen relacionados con la suavidad, aspereza, regularidad etc. Por lo que se aplican diversos métodos como la matriz de co-ocurrencia de nivel de gris (GLCM, de sus siglas en inglés), la cual busca cuantificar la textura de una imagen calculando las frecuencias del nivel de gris de pares de píxeles a diferentes distancias y ángulos.

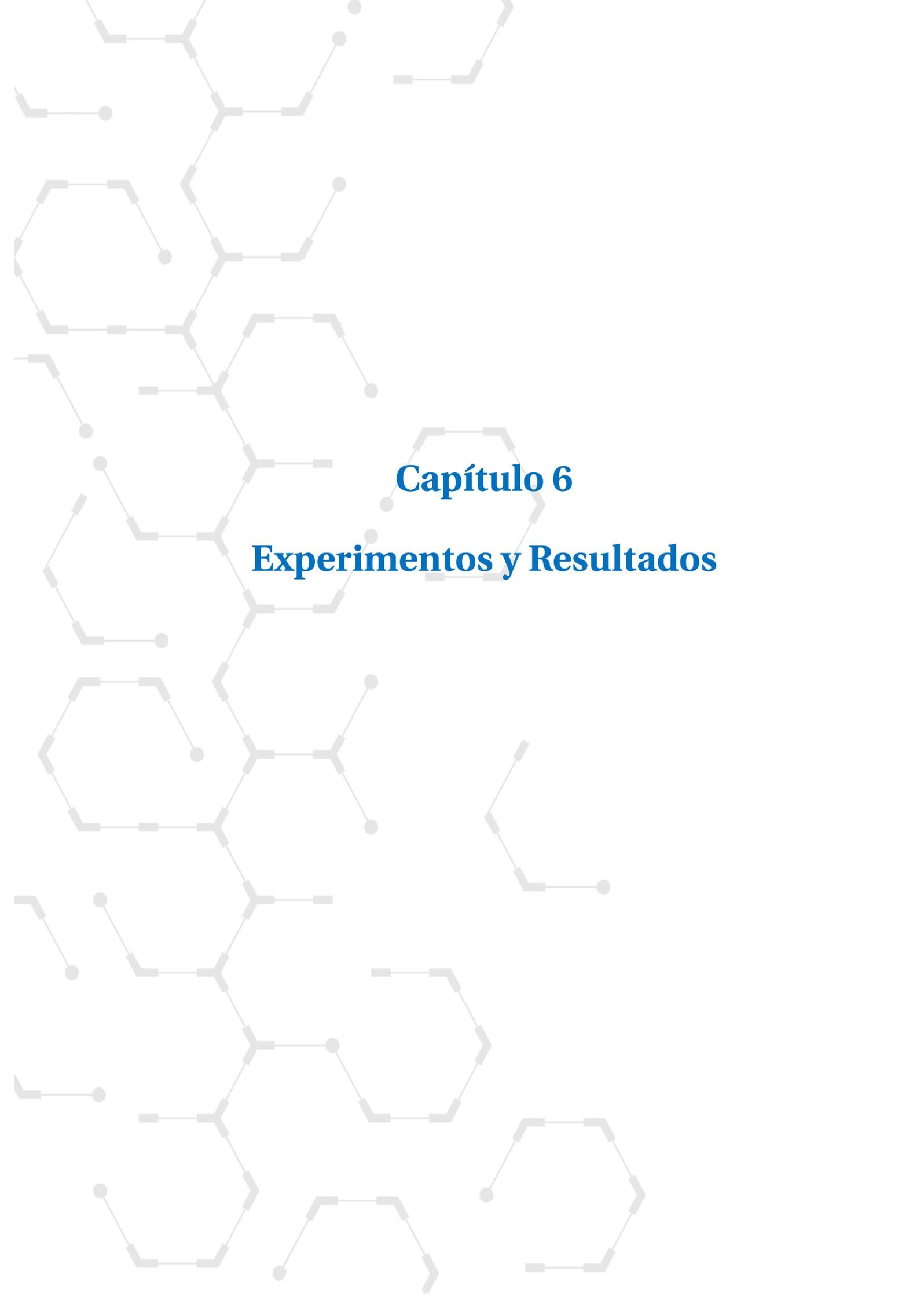
El siguiente tipo de características toma de base las bandas de la imagen transformadas utilizando el filtro denominado patrones binarios locales (LBP, de sus siglas en inglés), con lo que se busca describir la textura de la imagen analizando a los píxeles que rodean a cada píxel en lo individual. Con las imágenes filtradas con LBP procedimos a realizar el mismo análisis de textura derivado de GLCM con lo que generamos el

mismo número de características del paso anterior pero ahora con el filtro mencionado.

El último tipo de análisis de textura utilizado en este trabajo está basado en los filtros de Gabor. Para aplicar esta técnica generamos un banco de filtros haciendo variaciones de cinco frecuencias y ocho orientaciones lo que dio como resultado un banco de 40 filtros de Gabor. A partir de lo anterior, se aplicaron los 40 filtros a cada banda de las imágenes originales y se caracterizaron estadísticamente para generar los atributos que se plasmaron en la tabla de predictores.

Posteriormente se calculó el último tipo de características que busca cuantificar los atributos más sobresalientes que pueden identificarse en los bordes y contornos de las imágenes, para lo que utilizamos el análisis de momentos de Hu. Los cuales se basan en los momentos centrales de cada imagen y que son invariantes a transformaciones como traslación, escala y rotación. En total se calcularon 180,132 características para las imágenes Landsat y 185,430 para las imágenes Sentinel-2. La tabla de características resultante puede ser usada en algoritmos de clasificación supervisada.

Debido a la gran cantidad de características generadas se aplicó un método de selección de características con la finalidad de identificar a las más útiles e informativas para el problema de clasificación. El método utilizado se denomina Very Large Scale Relief (VLSRelief), con el cual se realizaron una serie de particiones secuenciales que se estudian en el siguiente capítulo.



Capítulo 6

Experimentos y Resultados

Capítulo 6 Experimentos y Resultados

6.1. Experimentos

En esta tesis se utilizaron dos conjuntos de datos de imágenes, Landsat correspondiente al territorio mexicano e imágenes de Sentinel-2 correspondiente al territorio de Europa. En este capítulo se describen los experimentos realizados con dichos conjuntos y se muestran los resultados.

Como se mencionó anteriormente, se generó un ordenamiento de características utilizando el método Very Large Scale ReliefF [22] (VLSReliefF). El mejor subconjunto obtenido se utilizó para alimentar a varios clasificadores.

Tratando de obtener el mejor resultado de clasificación posible, se compararon cuatro clasificadores muy utilizados en la literatura de percepción remota: RF, SVM, ET y ANN. Estos clasificadores se aplicaron a ambos conjuntos de datos y los resultados se describen más adelante.

Para los experimentos, los conjuntos de datos se dividieron en subconjuntos de entrenamiento y prueba. Con el 70% de entrenamiento se calculó el ranking de características, y con su resultado se entrenó cada clasificador lo que significa que obtenemos un resultado con un 10% de las mejores características para cada clasificador. Todas las exactitudes presentadas se calcularon con el 30% restante reservado solo para pruebas.

Como medida de rendimiento, se utilizó la exactitud general que se calcula como muestra la Ec. 6.1. Sabemos de las debilidades de esta métrica cuando los datos no están balanceados además de que en el caso de la base de datos de Landsat se buscó generar una BD balanceada también para subsanar ese hecho mostramos resultados por clase.

$$\text{exactitud} = \frac{\text{clasificados correctamente}}{\text{total de elementos}} \quad (6.1)$$

Aquí, la fórmula se refiere a una medida de la proporción de muestras clasificadas correctamente con respecto al total de elementos a clasificar. Aunque conocemos las desventajas de esta métrica, se seleccionó para fines de comparación porque la mayoría de los trabajos comparados la usaron.

Como se mencionó anteriormente, todas las características se han agrupado en 5 conjuntos según el tipo de información que representaban. Como resultado del procedimiento de selección de características, se utilizó el 10 % de las mejores características en conjunto con los cuatro clasificadores para obtener el rendimiento de exactitud. Además, con fines de comparación se utilizó el 100 % de las características en los procedimientos de evaluación de exactitud, es decir, sin tener en cuenta ninguna selección de características. Decidimos mostrar los resultados usando el 10 % de características porque con él se lograron los mejores resultados (en comparación con el 20 %, 40 %, etc.).

6.2. Resultados

Los resultados de las imágenes Landsat se muestran en la Tabla 6.1. Aquí, se puede ver que la mejor exactitud fue obtenida por el clasificador Extra Trees usando solo el 10 % de las mejores características y usando todos los grupos (marcadas en negritas). Sin embargo, utilizando solo características basadas en el grupo de Gabor se obtuvo un 76.6 % de exactitud, lo que también es un buen resultado. Comparando los resultados de los clasificadores usando todos los grupos de características filtradas solamente por el método de selección usando el mejor 10 % contra todas las características sin seleccionar es decir el 100 %; se puede decir que en la mayoría de los casos se obtuvieron los mejores valores de exactitud usando 10 %, solo con el clasificador SVM es mejor usar el 100 % de las características.

Con el propósito de conocer con más detalle qué características son más relevantes que otras, en la Tabla 6.2 se muestra el número de características de cada categoría dentro del mejor 10 % clasificadas, y también se muestra el número total de características en cada conjunto. De esta manera, podemos ver si el porcentaje con

Tabla 6.1: Desempeño de Landsat comparando el 10 % de las mejores características vs Todas.

Clasificador	Mejor 10% de Características						Todas las Características					
	STAT	GLCM	LBP	GABOR	HU	ALL	STAT	GLCM	LBP	GABOR	HU	ALL
RF	75.9%	66.9%	57.7%	76.9%	71.4%	76.8%	76.2%	67.8%	58.1%	75.6%	72.7%	75.7%
SVM	70.6%	66.8%	54.2%	65.2%	64.2%	66.2%	64.2%	65.6%	57.0%	70.4%	59.0%	72.1%
ET	74.3%	64.9%	56.3%	76.6%	68.4%	77.1%	75.2%	65.5%	55.8%	76.2%	70.0%	75.9%
ANN	63.7%	67.1%	56.2%	72.0%	63.7%	73.7%	69.0%	64.6%	55.6%	69.0%	62.7%	68.6%

respecto al total de características del grupo STAT, es alto o no en este 10%. También, podemos observar que el conjunto GABOR tiene una gran cantidad de características (15,330) en este conjunto del 10%, esto representa el 13% del total del mismo conjunto GABOR. Por el contrario, el conjunto GLCM tiene 1534 características en el mejor 10%, lo que representa menos del 5% de las características pertenecientes a este conjunto. Eso indica que algunos conjuntos tienen más presencia que otros en las primeras posiciones del ranking.

Además, la Tabla 6.2 muestra los tiempos de procesamiento para el mejor clasificador, en este caso para Landsat fue Extra Trees, usando el mejor 10% y todas las características para entrenamiento y prueba. Aquí, se puede observar que el uso de características de GABOR consume el mayor tiempo computacional: 3.4 ms por imagen en el entrenamiento y 0.50 ms en la prueba, estos tiempos corresponden solo para el paso de clasificación.

Para la extracción de características, en Landsat toma 6.58 segundos por imagen. En base a estos tiempos, está claro que si una gran cantidad de imágenes debe ser procesada y clasificada en tiempo real, considerando en tiempo real entre 20 y 30 imágenes por segundo, esto será imposible. Pero, en el caso práctico no obtenemos esta cantidad de imágenes satelitales por segundo además, nuestra comparación práctica es con el INEGI (Instituto Nacional de Estadística y Geografía de México) donde se realiza una clasificación de uso y cobertura de suelo para todo el territorio mexicano cada tres años aproximadamente porque utilizan un procedimiento predominantemente manual, y con esta metodología aplicada a su procedimiento como se indica en [43, 45], se podría generar una clasificación LULC nacional al menos una vez al año, la cual no reemplazaría el producto altamente detallado en número de clases, pero

podría contribuir a detectar cambios en clases mas generales pero con mayor oportunidad.

Tabla 6.2: Distribución de las características y tiempos de procesamiento con imagenes Landsat

Métrica	Mejor 10% de características ET						Todas las características ET					
	STAT	GLCM	LBP	GABOR	HU	ALL	STAT	GLCM	LBP	GABOR	HU	ALL
Características	224	1534	806	15330	119	18013	1734	31212	32232	114240	714	180132
Extracción (ms)	11.4	59.2	30.8	544	0.3	645.7	88	1205	1230	4054	2	6579
Entrenamiento (ms)	0.10	0.40	0.26	3.43	0.04	4.29	0.86	6.86	6.21	29.04	0.17	40.83
Prueba (ms)	0.07	0.07	0.07	0.50	0.07	0.67	0.07	1.23	1.07	4.23	0.07	10.10

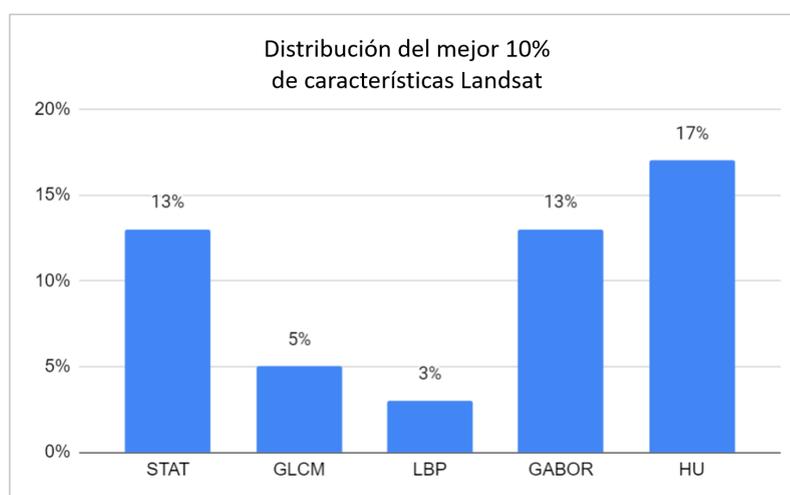


Figura 6.1: Histograma de distribución del mejor 10% de características Landsat. Cada barra indica el porcentaje relativo al total de características de cada grupo por separado.

La figura 6.1 muestra un diagrama de barras con la distribución del porcentaje de características del mejor 10%. Aquí, se puede ver que las características Estadísticas, Momentos de Hu y GABOR son las más relevantes usando imágenes Landsat. El porcentaje mostrado es relativo al total de las características de cada grupo en lo individual.

Buscando un análisis más profundo y comprensión de nuestros resultados, en la Tabla 6.3 se muestra el rendimiento por cada clase. Como se puede observar, la categoría mejor clasificada fue la de Cuerpo de Agua con un 96.6% de exactitud, y la peor clasificada fue Bosque Templado Mixto con un 58.4% de exactitud, lo que tiene sentido porque las tres primeras clases están relacionadas con Bosques Templados y

es difícil distinguir entre ellos.

Tabla 6.3: Exactitud por clase para el conjunto de datos Landsat, los resultados se alcanzaron con el algoritmo Extra-Trees.

Clase	Exactitud
1 Bosque templado de coníferas	73.6%
2 Bosque templado caducifolio	61.8%
3 Bosque templado mixto	58.4%
4 Bosque tropical perennifolio	81.3%
5 Bosque tropical caducifolio	73.7%
6 Matorral	73.7%
7 Vegetación de humedales	88.1%
8 Agricultura	70.8%
9 Pastizal	61.8%
10 Cuerpo de agua	96.6%
11 Tierra estéril	90.3%
12 Área urbana	93.3%

Para más detalles, en la Tabla 6.4 se presenta la matriz de confusión para las imágenes Landsat. Aquí, se puede observar qué clase es confundida por otras, y en las primeras filas de la matriz de confusión se puede ver, como se dijo antes, que las tres clases de bosque templado están altamente mezcladas lo que produce dificultades en su correcta clasificación.

Tabla 6.4: Matriz de confusión de los resultados de la clasificación para Landsat, los resultados se alcanzaron con el algoritmo Extra-Trees.

	Clase predicha											
	1	2	3	4	5	6	7	8	9	10	11	12
1	1084	87	205	27	15	17	6	11	12	0	0	0
2	60	993	204	20	133	16	2	17	67	0	0	0
3	277	225	810	54	34	45	4	28	29	1	0	3
4	22	13	35	1293	67	0	55	8	25	0	3	1
5	5	152	33	78	1132	13	3	52	58	0	1	2
6	4	3	40	0	24	1236	2	31	124	2	44	16
7	1	1	0	16	1	9	1386	23	25	15	47	5
8	11	22	20	51	48	56	18	1064	129	2	4	28
9	6	110	36	42	78	229	48	159	781	2	7	10
10	0	0	0	2	1	2	23	25	1	1406	30	1
11	0	0	1	1	0	40	22	11	6	27	1307	33
12	2	0	4	6	3	13	5	74	7	1	5	1388

Para propósitos de comparación, en la Tabla 6.5 se muestra el trabajo relacionado que usa el mismo tipo de datos para el Territorio Mexicano y el mismo número de clases. Nuestra metodología superó los resultados reportados en [28].

Tabla 6.5: Exactitud general de Landsat en comparación con el trabajo relacionado.

Trabajo	Exactitud general
MAD-MEX, [28]	71 %
La metodología propuesta	77.1 %

En el caso de las imágenes Sentinel-2, los resultados obtenidos se muestran en la Tabla 6.6. Como en el caso anterior con Landsat, el mejor resultado se logró utilizando solo el 10% de las mejores características con un 96.7% de exactitud. Sin embargo, también se obtienen buenos resultados utilizando todas las características.

Tabla 6.6: Desempeño de Sentinel-2 comparando el 10% de las mejores características vs Todas

Clasificador	Mejor 10% de Características						Todas las Características					
	STAT	GLCM	LBP	GABOR	HU	ALL	STAT	GLCM	LBP	GABOR	HU	ALL
RF	91.5%	94.7%	93.4%	94.2%	89.2%	95.3%	94.0%	94.7%	93.6%	95.0%	91.0%	95.7%
SVM	87.1%	95.2%	94.5%	92.9%	85.6%	95.2%	90.6%	94.1%	92.1%	93.2%	86.9%	94.5%
ET	90.9%	94.6%	93.2%	94.0%	87.6%	95.2%	93.3%	95.1%	93.4%	94.9%	89.6%	95.7%
ANN	92.1%	96.7%	95.5%	94.1%	90.0%	95.9%	94.1%	94.5%	93.9%	94.5%	89.5%	95.4%

La tabla 6.7 muestra el número de características de cada categoría en el 10% seleccionado respecto al número total que se muestra en la sección etiquetada con el 100%. El cálculo de los tiempos se realiza utilizando el clasificador ANN que fue el que mejores resultados mostró en los experimentos de clasificación. Al igual que Landsat, GABOR y momentos de Hu, son las grupos que están más presentes en las características que mostraron mejores resultados en la clasificación. A diferencia de Landsat, las características de LBP y GLCM ganan más relevancia con las imágenes de Sentinel-2. Esto se debe a que las imágenes Sentinel-2 tienen más resolución y capturan más información de la superficie terrestre, que se caracteriza mejor con descripciones de texturas como LBP y GLCM.

Tabla 6.7: Distribución de las características y tiempos de procesamiento con imágenes Sentinel-2

Métrica	Mejor 10% de características						Todas las características					
	STAT	GLCM	LBP	GABOR	HU	ALL	STAT	GLCM	LBP	GABOR	HU	ALL
Características	77	2388	3039	12948	91	18543	1785	32130	33180	117600	735	185430
Extracción (ms)	9.6	198.4	266.6	1224.6	0.6	1699.7	222.2	2668.9	2911.1	11122.2	4.4	16929
Entrenamiento (ms)	5.94	5.62	10.22	65.11	5.33	42.89	5.71	66.67	69.84	343.52	3.59	467.68
Prueba (ms)	0.15	0.15	0.15	0.59	0.15	0.89	0.15	1.33	1.26	5.56	0.15	8.67

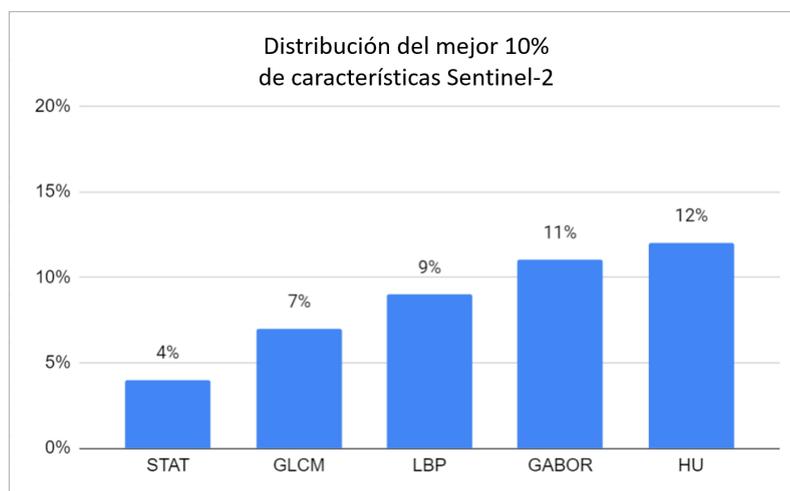


Figura 6.2: Histograma de distribución del mejor 10% de características Sentinel.

La figura 6.2 muestra la distribución de las categorías de las características dentro del 10% mejor clasificadas utilizando imágenes de Sentinel-2. Aquí, se puede observar que las características menos relevantes son la categoría STAT. Por el contrario los Momentos de HU y las características de Gabor son las más relevantes usando imágenes Sentinel-2.

La tabla 6.8 muestra las exactitudes logradas con el conjunto de datos de Sentinel-2 en cada clase. Aquí se pueden observar las mejores clases clasificadas, como Mar y Lago, Bosque y Residencial, con 99.6%, 98.8% y 98.5%, respectivamente. Como en resultados anteriores con Landsat, la categoría relacionada con el agua también es la mejor clasificada.

La matriz de confusión para el conjunto de datos de Sentinel-2 se presenta en la Tabla 6.9. Los resultados en Sentinel-2 son mejores que Landsat, por lo que la matriz de confusión concentra más los valores en su diagonal.

A modo de comparación, en la Tabla 6.10 se muestran todos los trabajos relacionados usando el mismo conjunto de datos de Sentinel-2 publicado originalmente en [36]. El mejor resultado reportado lo lograron los propios autores del conjunto de datos Sentinel-2, sin embargo, se han obtenido buenos resultados en otros trabajos. Tal es el caso de nuestra metodología que alcanza 96.7% de exactitud seguida por el resultado obtenido por los autores J. J. Senecal et al. reportada en su trabajo [83], donde

Tabla 6.8: Exactitud por clase para el conjunto de datos Sentinel-2, resultados alcanzados con el algoritmo ANN.

	Nombre	Exactitud
1	Cultivo anual	97.0 %
2	Bosque	98.8 %
3	Vegetación herbácea	95.9 %
4	Autopista	91.3 %
5	Industrial	96.9 %
6	Pastizal	94.4 %
7	Cultivo permanente	95.2 %
8	Residencial	98.5 %
9	Rio	97.8 %
10	Mar y Lago	99.6 %

Tabla 6.9: Matriz de confusión de los resultados de la clasificación para Sentinel-2, resultados alcanzados con el algoritmo ANN.

		Clase predicha									
		1	2	3	4	5	6	7	8	9	10
Clase etiquetada	1	863	1	4	10	0	8	12	0	0	0
	2	2	857	1	0	0	2	0	1	0	0
	3	2	2	859	3	1	4	10	2	0	0
	4	10	3	6	692	18	5	3	8	11	0
	5	0	0	1	19	724	0	2	2	1	0
	6	2	4	4	7	1	574	7	0	3	0
	7	11	0	21	14	0	11	690	1	0	0
	8	0	0	0	5	2	0	1	901	0	0
	9	0	0	0	8	1	4	0	0	750	4
	10	0	0	0	0	0	0	0	0	2	923

logran una exactitud de 96.6%.

Tabla 6.10: Exactitud general de Sentinel-2 en comparación con el trabajo relacionado

Trabajo	Exactitud General
E. Kroupi et al. [59]	87.6 %
S. Roy et al. [82]	93.2 %
G. Cavallaro et al. [14]	93.6 %
J. J. Senecal et al. [83]	96.6 %
La metodología propuesta	96.7 %
P. Helber et al. [36]	98.5 %

6.3. Reporte de variación en los experimentos

6.3.1. Variación en la exactitud

Para medir el impacto de la variación en la exactitud en el desempeño de los clasificadores se llevaron a cabo 1,000 experimentos, donde se fijó la muestra de entrenamiento y prueba. También se fijaron los hiperparámetros de los clasificadores que lograron los mejores resultados, Extra-Trees para Landsat y ANN para Sentinel. Y se utilizó el 10% de mejores características ordenadas usando el método de selección VLSReliefF. Con dichos experimentos buscamos cuantificar la variabilidad en el desempeño de los clasificadores.

En el caso de Landsat el clasificador utilizado fue Extra-Trees debido a que fue el que mejores resultados dio en la medida de exactitud utilizada para compararlo con otros clasificadores. Los hiperparámetros se dejaron con los valores por defecto, establecidos por la biblioteca Scikit-learn [74] por ejemplo: número de estimadores igual a 100, mínimo número de muestras para separar las ramas (`min_samples_split`) igual a 2, mínimo número de muestras por hoja (`min_samples_leaf`) igual a 1, por mencionar algunos. A excepción de los siguientes parámetros, los cuales se establecieron de la siguiente manera:

- Entropía como medida de pureza para la segmentación de los árboles.
- Se usó el 5% de las características para considerar la mejor segmentación.

El resultado de los experimentos se puede ver en el histograma de la Figura 6.3, donde se muestra la frecuencia de resultados en exactitud de los 1,000 experimentos realizados. El valor reportado fue 77.1% el cual se logró o superó 215 veces durante los experimentos. Se obtuvo un valor promedio de 77% de exactitud con una desviación estándar de 0.1%, un mínimo de 76.6% y un máximo de 77.5%.

En el caso de la base de datos de imágenes Sentinel-2 también se realizaron 1,000 experimentos donde el clasificador utilizado fue un Perceptrón Multicapa (MLP, por sus siglas en inglés de Multi Layer Perceptron). Utilizamos MLP debido a que fue

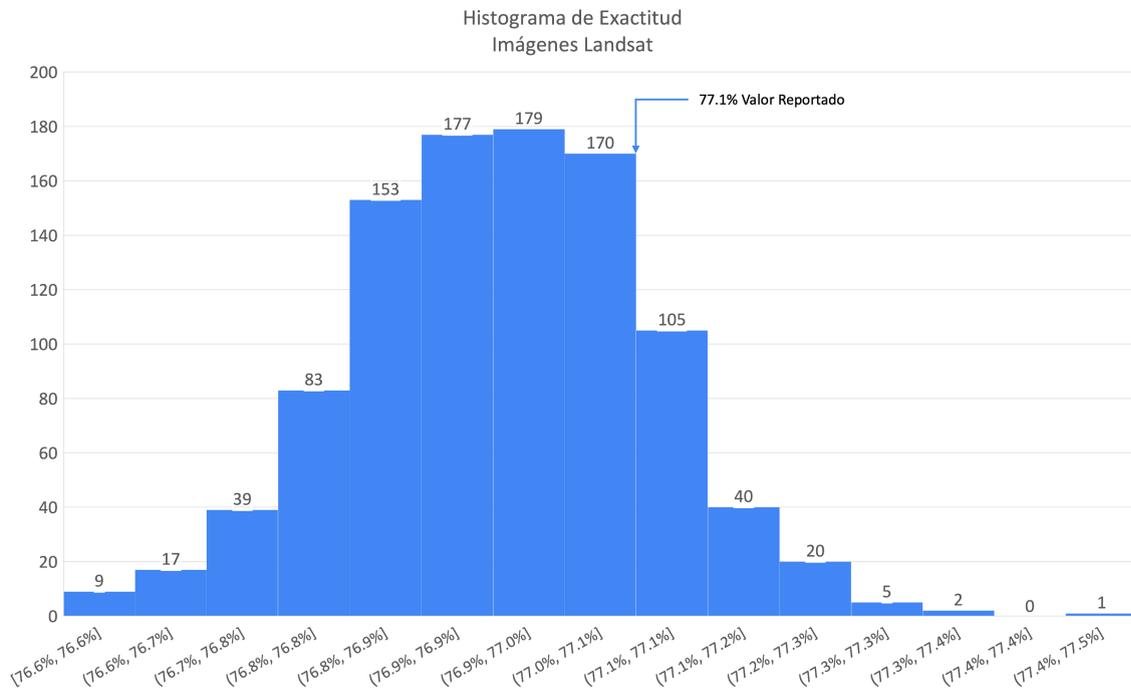


Figura 6.3: Histograma de la variación en la exactitud del clasificador Extra-Trees con muestras fijas para imágenes Landsat.

el clasificador que mejores resultados logró en comparación con los otros. Se usaron los hiperparámetros por defecto, establecidos por la biblioteca Scikit-learn [74], a excepción de el número máximo de iteraciones que se fijó en 300. Los resultados de los experimentos se pueden ver en la Figura 6.4. El valor reportado fue 96.7%, el cual se alcanzó o superó 214 veces en los experimentos. El valor promedio encontrado fue de 96.4% con una desviación estándar de 0.3%, un valor mínimo de 93.8% y un valor máximo de 97%.

6.3.2. Selección de características

En el Capítulo 1.6 se revisó el algoritmo de selección de características denominado VLSReliefF. Para ambos conjuntos de datos, Landsat y Sentinel-2, se aplicó dicho algoritmo seleccionando 350 subconjuntos de características de 350 atributos cada uno. A partir de lo cual se generaron 350 muestras de 350 características usando el 80% del conjunto de entrenamiento de cada base de datos de imágenes. En cada una de las iteraciones se realiza una medición de la calidad de cada atributo a partir de una evaluación iterativa usando el algoritmo ReliefF, explicado también en el Capítulo

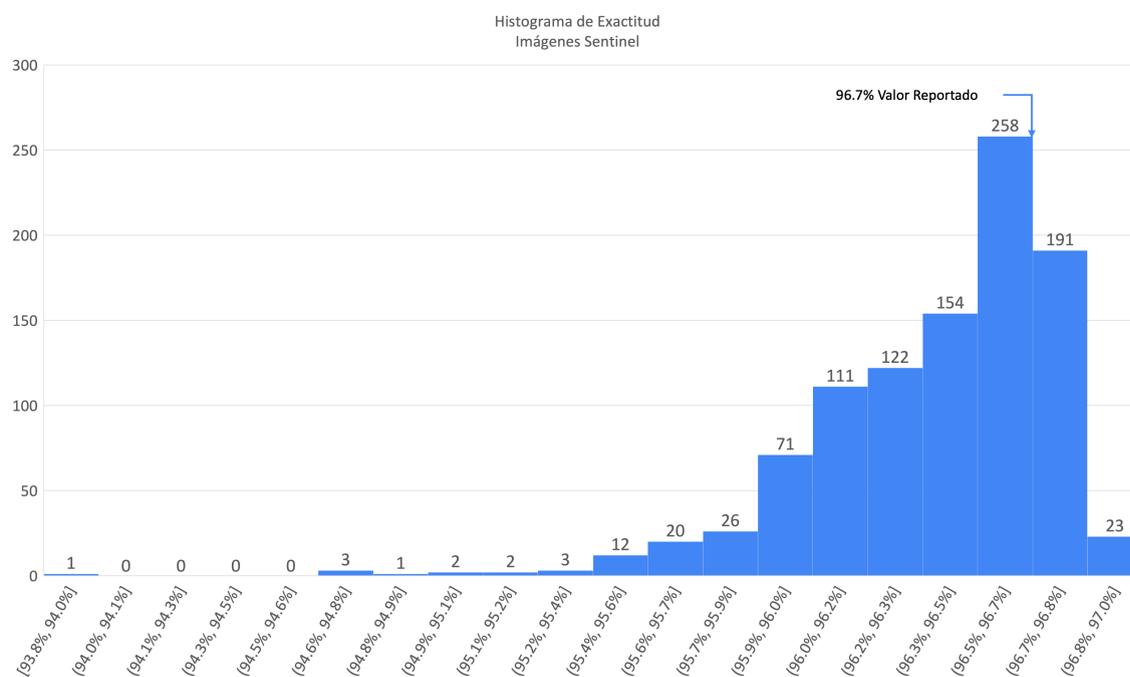


Figura 6.4: Histograma de la variación en la exactitud del clasificador MLP con muestras fijas para imágenes Sentinel.

1.6. Las estimaciones finales se calculan reemplazando los mejores resultados logrados para cada atributo individualmente en cada iteración, si estos son mayores a los encontrados en iteraciones anteriores. Al final de todas las iteraciones, que tardaron 72 horas en realizarse, se obtiene el ordenamiento de las características identificando aquellas que lograron los máximos valores para las estimaciones de calidad.

En la Figura 6.5 se muestra una representación gráfica del resultado del proceso de selección de características. El eje X representa la medición de la estimación de calidad lograda por el método VLSReliefF y el eje Y muestra una lista de las características ordenadas en función de su estimación de calidad correspondiente. Debido al gran número de características solo se muestran algunas de ellas. En la gráfica se agregó al nombre de cada característica el número correspondiente al orden que logró en el ordenamiento. Por lo tanto en la parte inferior de la gráfica se tienen las características con los mejores resultados en la estimación de calidad y en la parte superior se ven aquellas con los peores resultados.

Del lado izquierdo 1) se muestra la gráfica de Landsat cuyas características más significativas logran una estimación de 0.16. Del lado derecho 2) las características

Sentinel-2 con mayores estimaciones logran un 0.4. En ambos casos las peores estimaciones son negativas, lo cual indica que existen atributos que no contribuyen a la separación de clases. Se puede observar que las características de Sentinel logran un mejor discernimiento entre clases y, debido a que ambos conjuntos de características siguen la misma metodología de extracción, podemos pensar en que la diferencia se puede deber a la mayor resolución de las imágenes y a que las clases estén más claramente diferenciadas en el conjunto de datos de Sentinel. Usando el ordenamiento resultante se exploró la exactitud de forma incremental en ambos conjuntos de datos. El resultado coincidió en los dos casos, logrando el mejor valor de exactitud cuando se usó el mejor 10% de los atributos ordenados. Fue ese 10% el que se usó en los resultados reportados en la publicación arbitrada.

6.3.3. Variabilidad de las muestras de imágenes

Para explorar la variabilidad de las muestras de imágenes, se realizaron 1,000 experimentos con el clasificador SVM con kernel lineal con todos los parámetros por defecto, establecidos por la biblioteca Scikit-learn [74]. Se eligió dicho clasificador dado que se considera que la aleatoriedad que se incorpora en la clasificación es mínima y ésta depende de los datos de entrenamiento y prueba. A diferencia de otros clasificadores, Random Forest y Extra-Trees, que incorporan variabilidad aleatoria como parte de su estrategia de solución. Debido a que la aleatoriedad del clasificador SVM proviene de los datos de entrenamiento y prueba entonces esperamos observar es una medida de la variabilidad que cada conjunto de datos incorpora al problema de clasificación.

En los experimentos se utilizó solamente el 1% de las características más significativas con la finalidad de hacerlos viables computacionalmente.

Cada punto en la Figura 6.6 representa un experimento realizado con muestreo con reemplazo del 15% del total de la muestra de entrenamiento y evaluado con la muestra de prueba fija del 30% del total de cada conjunto de datos. En azul se pueden ver 1,000 experimentos con muestras aleatorias de tamaño fijo del 15% de las imágenes Sentinel-2 y en naranja los 1,000 experimentos con el mismo esquema de muestreo

variabilidad, lo cual puede deberse a dos factores, al menor número de clases (por lo que el traslape es menor) así como también a la mayor resolución espacial de las imágenes.

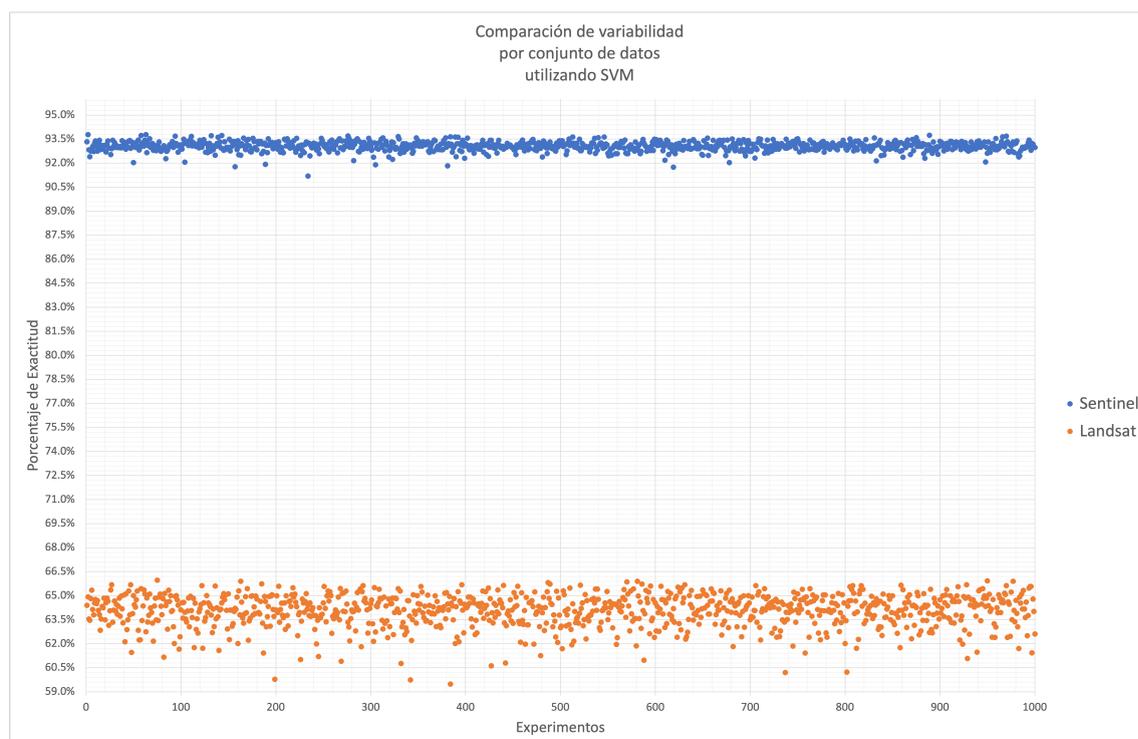


Figura 6.6: Comparación de la variación de los conjuntos de datos a través de los resultados obtenidos en 1,000 experimentos con el clasificador SVM. El eje X indica el número del experimento y el eje Y la exactitud lograda en los experimentos.

6.4. Resumen

En este capítulo se realizaron múltiples experimentos a partir de las bases de datos de entrenamiento generadas en el capítulo anterior. Para dichos experimentos los conjuntos de datos se dividieron en subconjuntos de entrenamiento y prueba, en una proporción 70%, 30% respectivamente. La métrica de desempeño utilizada fue la exactitud la cual es la proporción elementos clasificados correctamente respecto al total de elementos de la base de datos de prueba. Es importante señalar que el conjunto de prueba solo se utilizó una vez en los resultados reportados. Posteriormente, se analizó la variación en la exactitud en el desempeño de los clasificadores, para esto se llevaron a cabo 1000 evaluaciones a partir de muestras aleatorias del conjunto de

entrenamiento y se calculó la exactitud al evaluar contra el conjunto de prueba.

En el caso de las imágenes Landsat, los mejores resultados se observaron en los experimentos realizados con el 10% de las mejores variables predictoras según el algoritmo de selección de características, comparando el desempeño contra el total de las variables.

La exactitud lograda con el 10% de las mejores características, usando el clasificador Extra-Trees, fue de 77.1%.

Se realizaron mediciones de la distribución de las variables de acuerdo al tipo de procedimiento de extracción y los tiempos de procesamiento necesarios utilizando exclusivamente el algoritmo Extra-Trees para el caso de Landsat, debido a que fue el modelo con el mejor desempeño.

Se observó que las características más relevantes en el proceso de clasificación correspondieron a las derivadas de los Momentos de Hu con un 17% y las de textura derivadas de los filtros de Gabor con un 13%, al igual que las derivadas de el análisis estadístico de las imágenes.

También se calculó la matriz de confusión y la exactitud por clase para el conjunto de datos Landsat, lo cual nos permite observar que los peores desempeños se encontraron en aquellas clases con más proporción de mezcla encontrada en el conjunto de entrenamiento. Sin embargo, comparado con el trabajo relacionado, el resultado es satisfactorio porque lo supera con más del 6%.

En el caso de las imágenes Sentinel-2, el mejor resultado se obtuvo también con el 10% mejor de características. Aunque en este caso cuando se usaron todas las características también se obtuvieron muy buenos resultados. El mejor resultado fue de 96.7% de exactitud cuando se aplicó el algoritmo de clasificación de ANN al subconjunto de características derivadas de la matriz de co-ocurrencia de nivel de gris (GLCM).

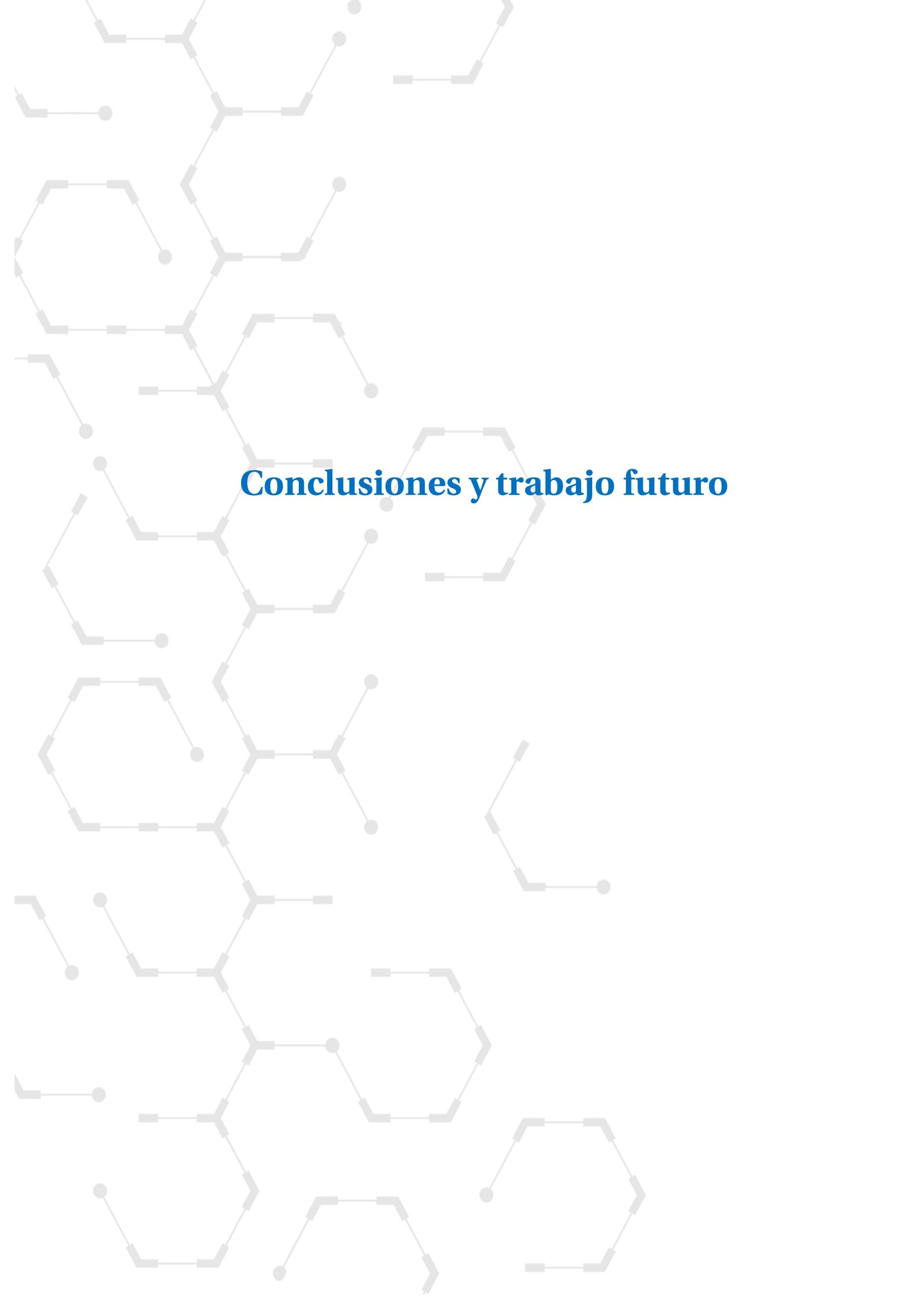
Al realizar el análisis de la distribución del mejor 10% de las características Sentinel-2 aplicado al algoritmo ganador, se logra observar que los momentos de Hu

está presente en un 12% y las características derivadas de los filtros de Gabor se encuentran en un 11%.

El resultado obtenido por la metodología propuesta logra un segundo lugar en comparación con 5 trabajos relacionados usando el mismo conjunto de datos.

En este capítulo se realizó una exploración de algunas fuentes de variación en los experimentos de clasificación. Se realizaron experimentos con las muestras de entrenamiento y validación fijas para cuantificar la variabilidad propia de los algoritmos de clasificación. También se muestran los resultados del algoritmo de selección de características dónde se puede ver un desempeño superior de las imágenes Sentinel-2.

Finalmente, se realizaron experimentos usando el algoritmo SVM debido a la menor aleatoriedad que incorpora al proceso de clasificación, con lo cual buscamos medir la variabilidad de los conjuntos de imágenes realizando un proceso de re-muestreo con reemplazo, se encontró una mayor variabilidad del conjunto de datos Landsat lo cual se podría deber a la menor resolución de las imágenes y a la mayor mezcla entre clases.



Conclusiones y trabajo futuro

Conclusiones y trabajo futuro

Las regiones de estudio de este trabajo son México y Europa para las cuales se utilizaron imágenes Landsat para cubrir México y Sentinel para Europa; aquí se presenta una metodología de extracción de características para la tarea de clasificación de Cobertura de Suelo y Uso de Suelo (LULC, por sus siglas en inglés, Land Use and Land Cover). Para esto, utilizamos imágenes calculadas que amplían el número de bandas originales mediante la generación de múltiples índices espectrales a nivel de píxel, generando 102 bandas para Landsat y 105 para Sentinel-2. Más tarde, se calcularon seis grupos de características, relacionadas con la forma, la textura y resúmenes estadísticos a partir de ellas.

Al final de este proceso de extracción, se generó un vector de características de muy alta dimensión para la descripción de la imagen, en busca de las características más informativas, se empleó un algoritmo de selección denominado VLSReliefF. Además, para la clasificación, se utilizaron varios clasificadores muy utilizados en la literatura de percepción remota (RF, SVM, ET y ANN), para comparar el desempeño en exactitud de los diferentes grupos de características. El conjunto de datos Landsat creado en este trabajo de investigación contiene 60,000 parches de imágenes Landsat etiquetados con 12 clases LULC, y está disponible para su uso y puede descargarse en Coronado y Moctezuma [16]. El resultado obtenido en este trabajo, 77.1% de exactitud, es una línea base que puede ser usada en otros trabajos de investigación relacionados con clasificación LULC.

En el caso de Sentinel-2, se logró un 96.7% de exactitud global. Los resultados obtenidos en este trabajo se compararon con cinco trabajos reportados recientemente en la literatura utilizando este mismo conjunto de datos. La metodología propuesta superó a cuatro de los cinco trabajos relacionados, y el restante tiene un resultado muy cercano. Como ventajas, la metodología propuesta es fácil de interpretar, y utilizando la estrategia de extracción propuesta, se podrían lograr resultados altamente competitivos.

Como resultado del análisis, se puede concluir que los grupos de características tienen diferentes grados de relevancia según el tipo de imagen a la que se aplican. Al utilizar Landsat, las características de textura más relevantes son derivadas de los filtros de Gabor. En el caso de Sentinel-2, las características derivadas de GLCM cobraron mayor relevancia. En ambos conjuntos de datos, también se lograron buenos resultados utilizando todas las características. Finalmente, también se puede observar que la relevancia de cada grupo de características está influenciada por la resolución de las imágenes a clasificar, siendo las características relacionadas con la textura y la forma, las más importantes en Sentinel-2 y en el caso de Landsat son especialmente relevantes los relacionados con resúmenes estadísticos.

Como trabajo futuro se considera aplicar la metodología desarrollada en el presente documento a otro tipo de clases, como la detección de crecimiento urbano o la delimitación y clasificación de la frontera agrícola. Así como la realización de mapas nacionales con las clasificaciones resultantes.

También se considera aplicar otras estrategias de segmentación, por ejemplo, en lugar de usar parches de imágenes regulares, explorar el impacto de aplicar la técnica de análisis de imágenes basadas en objetos (OBIA, de sus siglas en inglés) con agrupaciones irregulares de píxeles.

Por último, para lidiar con las clases que resultaron más difíciles de clasificar, como las autopistas, se considera explorar características basadas en palabras visuales a través del análisis no supervisado del conjunto de entrenamiento.

Bibliografía

- [1] *Guía para la interpretación de cartografía. Uso del suelo y vegetación. Escala 1:250, 000. Serie VI.* INEGI, 2017.
- [2] *Producción y publicación de la Geomediana Nacional a partir de imágenes del Cubo de Datos Geoespaciales de México: Documento metodológico.* INEGI, México, 2020.
- [3] ADAM, F., ESCH, T., AND DATCU, M. Feature investigation for large scale urban detection using landsat imagery. *Proceedings 2*, 7 (2018). [doi:10.3390/ecrs-2-05162].
- [4] ANDERSON, K., RYAN, B., SONNTAG, W., KAVVADA, A., AND FRIEDL, L. Earth observation in service of the 2030 agenda for sustainable development. *Geo-spatial Information Science 20*, 2 (2017), 77–96. [doi:10.1080/10095020.2017.1333230].
- [5] ARSET. Applied remote sensing training program. [url:<https://appliedsciences.nasa.gov/what-we-do/capacity-building/arset>].
- [6] BAIG, M. H. A., ZHANG, L., SHUAI, T., AND TONG, Q. Derivation of a tasselled cap transformation based on landsat 8 at-satellite reflectance. *Remote Sensing Letters 5*, 5 (2014), 423–431.
- [7] BELWARD, A. S., AND SKØIEN, J. O. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS Journal of Photogrammetry and Remote Sensing 103* (2015), 115–128. [doi:10.1016/j.isprsjprs.2014.03.009].
- [8] BISHOP, C. M. *Pattern recognition and machine learning.* springer, 2006.
- [9] BOUZEKRI, S., LASBET, A. A., AND LACHEHAB, A. A new spectral index for extraction of built-up area using landsat-8 data. *Journal of the Indian Society of Remote Sensing 43*, 4 (2015), 867–873.

- [10] BREIMAN, L. Random forests. *Machine learning* 45, 1 (2001), 5–32. [doi:10.1023/A:1010933404324].
- [11] BURKOV, A. The hundred-page machine learning book (2019). ISBN-13, 978–1999579500.
- [12] CAI, Y., GUAN, K., PENG, J., WANG, S., SEIFERT, C., WARDLOW, B., AND LI, Z. A high-performance and in-season classification system of field-level crop types using time-series landsat data and a machine learning approach. *Remote Sensing of Environment* 210, Agric. For. Meteorol. 173 2013 (6 2018), 35–47. [doi:10.1016/j.rse.2018.02.045].
- [13] CAMPBELL, J. B., AND WYNNE, R. H. *Introduction to remote sensing*. Guilford Press, 2011.
- [14] CAVALLARO, G., BAZI, Y., MELGANI, F., AND RIEDEL, M. Multi-scale convolutional svm networks for multi-class classification problems of remote sensing images. In *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium* (July 2019), pp. 875–878. [doi:10.1109/IGARSS.2019.8899831].
- [15] CONABIO. Monitoreo de la cobertura de suelo.
- [16] CORONADO, A., AND MOCTEZUMA, D. Multispectral and augmented Landsat data with land cover labels, June 2020. [doi:10.5281/zenodo.3891580].
- [17] CUI, J., ZHANG, X., WANG, W., AND WANG, L. Integration of optical and sar remote sensing images for crop-type mapping based on a novel object-oriented feature selection method. *International Journal of Agricultural and Biological Engineering* 13, 1 (2020), 178–190.
- [18] DARIUS PHIRI, J. M. Developments in Landsat Land Cover Classification Methods: A Review. *Remote Sensing* 9, 9 (2017), 967.
- [19] DELGADILLO-HERRERA, M., ARREOLA-ESQUIVEL, M., TOXQUI-QUITL, C., AND PADILLA-VIVANCO, A. Normalized difference indices in Landsat 5 TM satellite data. In *Current Developments in Lens Design and Optical Engineering XX* (2019),

- R. B. Johnson, V. N. Mahajan, and S. Thibault, Eds., vol. 11104, International Society for Optics and Photonics, SPIE, pp. 225 – 232. [doi:10.1117/12.2532322].
- [20] DENIS, G., CLAVERIE, A., PASCO, X., DARNIS, J.-P., MAUPEOU, B. D., LAFAYE, M., AND MOREL, E. Towards disruptions in Earth observation? New Earth Observation systems and markets evolution: Possible scenarios and impacts. *Acta Astronautica* 137 (2017), 415–433. [doi:10.1016/j.actaastro.2017.04.034].
- [21] DOS LUCIANO, A., PICOLI, M., ROCHA, J., FRANCO, H., SANCHES, G., LEAL, M., AND LE MAIRE, G. Generalized space-time classifiers for monitoring sugarcane areas in Brazil. *Remote Sensing of Environment* 215 (2018), 438–451.
- [22] EPPSTEIN, M. J., AND HAAKE, P. Very Large Scale ReliefF for Genome-Wide Association Analysis. *2008 IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology* (2008), 112–119. [doi:10.1109/cibcb.2008.4675767].
- [23] FENG, D., YU, L., ZHAO, Y., CHENG, Y., XU, Y., LI, C., AND GONG, P. A multiple dataset approach for 30-m resolution land cover mapping: a case study of continental Africa. *International Journal of Remote Sensing* 39, 12 (2018), 3926–3938. [doi:10.1080/01431161.2018.1452073].
- [24] G AZZARI, D. L. Landsat-based classification in the cloud: An opportunity for a paradigm shift in land cover monitoring. *Remote Sensing of Environment* 202 (2017), 64–74.
- [25] GABOR, D. Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering* 93, 26 (1946), 429–441.
- [26] GAMON, J., AND SURFUS, J. Assessing leaf pigment content and activity with a reflectometer. *The New Phytologist* 143, 1 (1999), 105–117.
- [27] GAVIN, D., DHU, T., SAGAR, S., MUELLER, N., DUNN, B., LEWIS, A., LYMBURNER, L., MINCHIN, S., OLIVER, S., ROSS, J., AND THANKAPPAN, M. Digital earth australia - from satellite data to better decisions. In *IGARSS 2018 - 2018 IEEE*

- International Geoscience and Remote Sensing Symposium* (July 2018), pp. 8633–8635. [doi:10.1109/IGARSS.2018.8518160].
- [28] GEBHARDT, S., WEHRMANN, T., RUIZ, M., MAEDA, P., BISHOP, J., SCHRAMM, M., KOPEINIG, R., CARTUS, O., KELLNDORFER, J., RESSL, R., SANTOS, L., AND SCHMIDT, M. Mad-mex: Automatic wall-to-wall land cover monitoring for the mexican redd-mrv program using all landsat data. *Remote Sensing* 6, 5 (4 2014), 3923–3943. [doi:10.3390/rs6053923].
- [29] GEORGANOS, S., GRIPPA, T., VANHUYSSSE, S., LENNERT, M., SHIMONI, M., KALOGIROU, S., AND WOLFF, E. Less is more: optimizing classification performance through feature selection in a very-high-resolution remote sensing object-based urban application. *GIScience & Remote Sensing* 55, 2 (2017), 221–242. [doi:10.1080/15481603.2017.1408892].
- [30] GÉRON, A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O’Reilly Media, 2019.
- [31] GEURTS, P., ERNST, D., AND WEHENKEL, L. Extremely randomized trees. *Machine learning* 63, 1 (2006), 3–42. [doi:10.1007/s10994-006-6226-1].
- [32] GITELSON, A. A., GRITZ, Y., AND MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of plant physiology* 160, 3 (2003), 271–282.
- [33] GITELSON, A. A., KAUFMAN, Y. J., AND MERZLYAK, M. N. Use of a green channel in remote sensing of global vegetation from eos-modis. *Remote sensing of Environment* 58, 3 (1996), 289–298.
- [34] GONZALEZ, R., AND WOODS, R. *Digital image processing: Pearson prentice hall. Upper Saddle River, NJ 1* (2008).
- [35] HARALICK, R. M., SHANMUGAM, K., AND DINSTEN, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics SMC-3*, 6 (Nov 1973), 610–621. [doi:10.1109/TSMC.1973.4309314].

- [36] HELBER, P., BISCHKE, B., DENGEL, A., AND BORTH, D. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* (2019). [doi:10.1109/JSTARS.2019.2918242].
- [37] HINTON, G. E. Connectionist learning procedures. In *Machine learning*. Elsevier, 1990, pp. 555–610.
- [38] HOLLOWAY, J., AND MENGERSEN, K. Statistical machine learning methods and remote sensing for sustainable development goals: A review. *Remote Sensing* 10, 9 (2018). [doi:10.3390/rs10091365].
- [39] HOSSAIN, M. A., JIA, X., AND BENEDIKTSSON, J. A. One-Class Oriented Feature Selection and Classification of Heterogeneous Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 9, 4 (2016), 1606–1612. [doi:10.1109/jstars.2015.2506268].
- [40] HUETE, A., DIDAN, K., MIURA, T., RODRIGUEZ, E., GAO, X., AND FERREIRA, L. Overview of the radiometric and biophysical performance of the modis vegetation indices. *Remote Sensing of Environment* 83, 1 (2002), 195 – 213. [doi:10.1016/S0034-4257(02)00096-2].
- [41] HUETE, A. R. A soil-adjusted vegetation index (savi). *Remote sensing of environment* 25, 3 (1988), 295–309.
- [42] INEGI. Uso del suelo y vegetación, escala 1:250000, serie V (continuo nacional). <http://geoportal.conabio.gob.mx/metadatos/doc/html/usv250s5ugw.html>, 2013. [Online; accessed 5-October-2020].
- [43] INEGI. Inegi presenta carta del uso del suelo y vegetación serie vi, 2017. https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2017/especiales/especiales2017_12_01.pdf.
- [44] INEGI. Uso de suelo y vegetación. <https://www.inegi.org.mx/temas/usosuelo/>, 2020. [Online; accessed 13-November-2020].

- [45] INEGI. Uso del suelo y vegetación serie vi, 2020. <https://www.inegi.org.mx/temas/usosuelo/default.html>.
- [46] IRELAND, G., VOLPI, M., AND PETROPOULOS, G. P. Examining the capability of supervised machine learning classifiers in extracting flooded areas from landsat tm imagery: A case study from a mediterranean flood. *Remote Sensing* 7, 3 (2015), 3372–3399. [doi:10.3390/rs70303372].
- [47] IZENMAN, A. J. Modern multivariate statistical techniques. *Regression, classification and manifold learning* 10 (2013), 978–0.
- [48] JAMES, G., WITTEN, D., HASTIE, T., AND TIBSHIRANI, R. *An introduction to statistical learning*, vol. 112. Springer, 2013.
- [49] JAMSRAN, B.-E., LIN, C., BYAMBAKHUU, I., RAASH, J., AND AKHMADI, K. Applying a support vector model to assess land cover changes in the uvs lake basin ecoregion in mongolia. *Information Processing in Agriculture* (2018).
- [50] JASON GILBERTSON, JACO KEMP, A. V. N. Effect of pan-sharpening multi-temporal Landsat 8 imagery for crop type differentiation using different classification techniques. *Computers and Electronics in Agriculture* 134 (2017), 151–159.
- [51] JIELI, C., MANCHUN, L., YONGXUE, L., CHENGLI, S., AND WEI, H. Extract residential areas automatically by new built-up index. In *2010 18th International Conference on Geoinformatics* (June 2010), pp. 1–5. [doi:10.1109/GEOINFORMATICS.2010.5567823].
- [52] JORDAN, C. F. Derivation of leaf-area index from quality of light on the forest floor. *Ecology* 50, 4 (1969), 663–666. [doi:10.2307/1936256].
- [53] KAUFMAN, Y. J., AND TANRE, D. Atmospherically resistant vegetation index (arvi) for eos-modis. *IEEE Transactions on Geoscience and Remote Sensing* 30, 2 (March 1992), 261–270. [doi:10.1109/36.134076].
- [54] KAWAMURA, M. Relation between social and environmental conditions in colombo sri lanka and the urban index estimated by satellite remote sensing data.

- In *Proc. 51st Annual Conference of the Japan Society of Civil Engineers* (1996), pp. 190–191.
- [55] KILLOUGH, B. Overview of the Open Data Cube Initiative. *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium 00* (2018), 8629–8632. [doi:10.1109/igarss.2018.8517694].
- [56] KIRA, K., AND RENDELL, L. A. A practical approach to feature selection. In *Machine learning proceedings 1992*. Elsevier, 1992, pp. 249–256.
- [57] KONONENKO, I. Estimating attributes: Analysis and extensions of relief. In *Machine Learning: ECML-94* (Berlin, Heidelberg, 1994), F. Bergadano and L. De Raedt, Eds., Springer Berlin Heidelberg, pp. 171–182. [doi:10.1007/3-540-57868-4_57].
- [58] KONONENKO, I. Estimating attributes: Analysis and extensions of relief. In *European conference on machine learning* (1994), Springer, pp. 171–182.
- [59] KROUPI, E., KESA, M., NAVARRO-SÁNCHEZ, V. D., SAEED, S., PELLOQUIN, C., ALHADDAD, B., MORENO, L., SORIA-FRISCH, A., AND RUFFINI, G. Deep convolutional neural networks for land-cover classification with Sentinel-2 images. *Journal of Applied Remote Sensing* 13, 2 (2019), 1 – 22. [doi:10.1117/1.JRS.13.024525].
- [60] KUSSUL, N., LAVRENIUK, M., KOLOTII, A., SKAKUN, S., RAKOID, O., AND SHUMILO, L. A workflow for sustainable development goals indicators assessment based on high-resolution satellite data. *International Journal of Digital Earth* 0, 0 (2019), 1–13. [doi:10.1080/17538947.2019.1610807].
- [61] KUSSUL, N., MYKOLA, L., SHELESTOV, A., AND SKAKUN, S. Crop inventory at regional scale in Ukraine: developing in season and end of season crop maps with multi-temporal optical and SAR satellite imagery. *European Journal of Remote Sensing* 51, 1 (2018), 627–636.
- [62] KWAK, G.-H., AND PARK, N.-W. Impact of texture information on crop classification with machine learning and uav images. *Applied Sciences* 9, 4 (2019). [doi:10.3390/app9040643].

- [63] LEAVITT, B. Remote estimation of leaf area index and green leaf biomass in maize canopies. *Geophys Res Lett* 30, 5 (2003), 52–152.
- [64] LEWIS, A., OLIVER, S., LYMBURNER, L., EVANS, B., WYBORN, L., MUELLER, N., RAEVKSI, G., HOOKE, J., WOODCOCK, R., SIXSMITH, J., WU, W., TAN, P., LI, F., KILLOUGH, B., MINCHIN, S., ROBERTS, D., AYERS, D., BALA, B., DWYER, J., DEKKER, A., DHU, T., HICKS, A., IP, A., PURSS, M., RICHARDS, C., SAGAR, S., TRENHAM, C., WANG, P., AND WANG, L.-W. The australian geoscience data cube — foundations and lessons learned. *Remote Sensing of Environment* 202 (2017), 276 – 292. [doi:10.1016/j.rse.2017.03.015].
- [65] LI, M., IM, J., AND BEIER, C. Machine learning approaches for forest classification and change analysis using multi-temporal landsat tm images over huntington wildlife forest. *GIScience & Remote Sensing* 50, 4 (2013), 361–384. [doi:10.1080/15481603.2013.819161].
- [66] LUOJIA HU, WENYU LI, B. X. Monitoring mangrove forest change in China from 1990 to 2015 using Landsat-derived spectral-temporal variability metrics. *International Journal of Applied Earth Observation and Geoinformation* 73 (2018), 88–98.
- [67] MCFEETERS, S. K. The use of the normalized difference water index (ndwi) in the delineation of open water features. *International journal of remote sensing* 17, 7 (1996), 1425–1432.
- [68] MING-KUEI HU. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* 8, 2 (February 1962), 179–187. [doi:10.1109/TIT.1962.1057692].
- [69] MOODY, D. I., BAUER, D. E., BRURNBY, S. P., CHISOLM, E. D., WARREN, M. S., SKILLMAN, S. W., AND KEISLER, R. Land Cover Classification in Fused Multisensor Multispectral Satellite Imagery. *2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)* (2016), 85–88. [doi:10.1109/ssiai.2016.7459181].

- [70] MOUNTRAKIS, G., IM, J., AND OGOLE, C. Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing* 66, 3 (2011), 247–259. [doi:10.1016/j.isprsjprs.2010.11.001].
- [71] NEDKOV, R. Orthogonal transformation of segmented images from the satellite sentinel-2. *Comptes rendus de l'Academie bulgare des Sciences* 70, 5 (2017), 687–692.
- [72] OJALA, T., PIETIKAINEN, M., AND MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 7 (July 2002), 971–987. [doi:10.1109/TPAMI.2002.1017623].
- [73] OUMA, Y. O. Advancements in medium and high resolution Earth observation for land-surface imaging: Evolutions, future trends and contributions to sustainable development. *Advances in Space Research* 57 (2016), 110–126. [doi:10.1016/j.asr.2015.10.038].
- [74] PEDREGOSA, F., VAROQUAUX, G., GRAMFORT, A., MICHEL, V., THIRION, B., GRISSEL, O., BLONDEL, M., PRETTENHOFER, P., WEISS, R., DUBOURG, V., ET AL. Scikit-learn: Machine learning in python. *the Journal of machine Learning research* 12 (2011), 2825–2830.
- [75] PHALKE, A. R., AND ÖZDOĞAN, M. Large area cropland extent mapping with landsat data and a generalized classifier. *Remote Sensing of Environment* 219, Sensors 16 9 2016 (12 2018), 180–195. [doi:10.1016/j.rse.2018.09.025].
- [76] PICOLI, M., CAMARA, G., SOND ROLF SIMÕES, I. S., CARVALHO, A., MACIEL, A., COUTINHO, A., ESQUERDO, J., ANTUNES, J., BEGOTTI, R., ARVOR, D., AND ALMEIDA, C. Big earth observation time series analysis for monitoring Brazilian agriculture. *ISPRS Journal of Photogrammetry and Remote Sensing* (2018).
- [77] ROBERTS, D., MUELLER, N., AND MCINTYRE, A. High-dimensional pixel composites from earth observation time series. *IEEE Transactions on Geoscience and Remote Sensing* 55, 11 (Nov 2017), 6254–6264. [doi:10.1109/TGRS.2017.2723896].

- [78] ROBNIK-ŠIKONJA, M., AND KONONENKO, I. Theoretical and empirical analysis of relief and relief. *Machine learning* 53, 1 (2003), 23–69.
- [79] RODRIGUEZ-GALIANO, V. F., AND CHICA-RIVAS, M. Evaluation of different machine learning methods for land cover mapping of a mediterranean area using multi-seasonal landsat images and digital terrain models. *International Journal of Digital Earth* 7, 6 (2014), 492–509. [doi:10.1080/17538947.2012.748848].
- [80] ROGERS, A. S., AND KEARNEY, M. S. Reducing signature variability in unmixing coastal marsh thematic mapper scenes using spectral indices. *International Journal of Remote Sensing* 25, 12 (2004), 2317–2335.
- [81] ROUSE, J. W., HAAS, R. H., SCHELL, J. A., DEERING, D. W., ET AL. Monitoring vegetation systems in the great plains with erts. *NASA Spec. Publ* 351, 1 (1974), 309.
- [82] ROY, S., SANGINETO, E., SEBE, N., AND DEMIR, B. Semantic-fusion gans for semi-supervised satellite image classification. In *2018 25th IEEE International Conference on Image Processing (ICIP)* (Oct 2018), pp. 684–688. [doi:10.1109/ICIP.2018.8451836].
- [83] SENEAL, J. J., SHEPPARD, J. W., AND SHAW, J. A. Efficient convolutional neural networks for multi-spectral image classification. In *2019 International Joint Conference on Neural Networks (IJCNN)* (July 2019), pp. 1–8. [doi:10.1109/IJCNN.2019.8851840].
- [84] SHAHARUM, N., SHAFRI, H., GAMBO, J., AND ABIDIN, F. Mapping of Krau Wildlife Reserve (KWR) Protected Area Using Landsat 8 and Supervised Classification Algorithms. *Remote Sensing Applications: Society and Environment* (2018).
- [85] SHAHRIAR S HEYDARI, G. M. Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 Landsat sites. *Remote Sensing of Environment* 204 (2018), 648–658.

- [86] SHEN, H., LIN, Y., TIAN, Q., XU, K., AND JIAO, J. A comparison of multiple classifier combinations using different voting-weights for remote sensing image classification. *International Journal of Remote Sensing* 39, 11 (2018), 3705–3722.
- [87] TEFFERA, Z. L., LI, J., DEBSU, T. M., AND MENEGESHA, B. Y. Assessing land use and land cover dynamics using composites of spectral indices and principal component analysis: A case study in middle awash subbasin, ethiopia. *Applied Geography* 96, Applied Geography 31 2011 (7 2018), 109–129. [doi:10.1016/j.apgeog.2018.05.015].
- [88] TELUGUNTLA, P., THENKABAIL, P. S., OLIPHANT, A., XIONG, J., GUMMA, M. K., CONGALTON, R. G., YADAV, K., AND HUETE, A. A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. *ISPRS Journal of Photogrammetry and Remote Sensing* 144 (2018), 325–340. [doi:10.1016/j.isprsjprs.2018.07.017].
- [89] TOTH, C., AND JÓZKÓW, G. Remote sensing platforms and sensors: A survey. *ISPRS Journal of Photogrammetry and Remote Sensing* 115 (2016), 22–36. [doi:10.1016/j.isprsjprs.2015.10.004].
- [90] UNION, E. Commission delegated regulation (eu) no 1159/2013 of 12 july 2013 supplementing regulation (eu) no 911/2010 of the european parliament and of the council on the european earth monitoring programme(gmes) by establishing registration and licensing conditions for gmes users and defining criteria for restricting access to gmes dedicated data and gmesservice information. *Official Journal of the European Union* (2013), 1–6. [url:https://eur-lex.europa.eu/eli/reg_del/2013/1159/oj].
- [91] URBANOWICZ, R. J., MEEKER, M., LA CAVA, W., OLSON, R. S., AND MOORE, J. H. Relief-based feature selection: Introduction and review. *Journal of Biomedical Informatics* 85 (2018), 189 – 203. [doi:10.1016/j.jbi.2018.07.014].
- [92] URBANOWICZ, R. J., OLSON, R. S., SCHMITT, P., MEEKER, M., AND MOORE, J. H. Benchmarking Relief-Based Feature Selection Methods for Bioinformatics Data

- Mining. [doi:10.1016/j.jbi.2018.07.015].
- [93] VICTOR F RODRIGUEZ-GALIANO, M. C.-R. Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. *International Journal of Digital Earth* 7, 6 (2012), 492–509.
- [94] WAGNER, M. A., MYINT, S. W., AND CERVENY, R. S. Geospatial assessment of recovery rates following a tornado disaster. *IEEE Transactions on Geoscience and Remote Sensing* 50, 11 (2012), 4313–4322.
- [95] WANG, Z., ZHANG, Y., CHEN, Z., YANG, H., SUN, Y., KANG, J., YANG, Y., AND LIANG, X. Application of Relieff Algorithm to Selecting Feature Sets for Classification of High Resolution Remote Sensing Image. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (2016), 755–758. [doi:10.1109/igarss.2016.7729190].
- [96] WANHUI CHEN, LIANGYUN LIU, CHAO ZHANG, JIHUA WANG, JINDI WANG, AND YUCHUN PAN. Monitoring the seasonal bare soil areas in beijing using multitemporal tm images. In *IGARSS 2004. 2004 IEEE International Geoscience and Remote Sensing Symposium* (Sep. 2004), vol. 5, pp. 3379–3382 vol.5. [doi:10.1109/36.134076].
- [97] WAQAR, M., MIRZA, J., MUMTAZ, R., AND HUSSAIN, E. Development of new indices for extraction of built-up area & bare soil from landsat data. *open access scientific reports*, 1 (1), 2012.
- [98] WILSON, E. H., AND SADER, S. A. Detection of forest harvest type using multiple dates of landsat tm imagery. *Remote Sensing of Environment* 80, 3 (2002), 385–396.
- [99] WOODCOCK, C. E., ALLEN, R., ANDERSON, M., BELWARD, A., BINDSCHADLER, R., COHEN, W., GAO, F., GOWARD, S. N., HELDER, D., HELMER, E., ET AL. Free access to landsat imagery. *Science* 320, 5879 (2008), 1011–1011.

- [100] XIONG, J., THENKABAIL, P., TILTON, J., GUMMA, M., TELUGUNTLA, P., OLIPHANT, A., CONGALTON, R., YADAV, K., AND GORELICK, N. Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine. *Remote Sensing* 9, 10 (2017), 1065. [doi:10.3390/rs9101065].
- [101] XU, H. Modification of normalised difference water index (ndwi) to enhance open water features in remotely sensed imagery. *International journal of remote sensing* 27, 14 (2006), 3025–3033.
- [102] XU, H. A new index for delineating built-up land features in satellite imagery. *International Journal of Remote Sensing* 29, 14 (2008), 4269–4276. [doi:10.1080/01431160802039957].
- [103] YE, M., XU, Y., JI, C., CHEN, H., LU, H., AND QIAN, Y. Feature selection for cross-scene hyperspectral image classification using cross-domain relieff. *International Journal of Wavelets, Multiresolution and Information Processing* 17, 05 (2019), 1950039.
- [104] ZHA, Y., GAO, J., AND NI, S. Use of normalized difference built-up index in automatically mapping urban areas from tm imagery. *International Journal of Remote Sensing* 24, 3 (2003), 583–594. [doi:10.1080/01431160304987].
- [105] ZHOU, Y., ZHANG, R., WANG, S., AND WANG, F. Feature Selection Method Based on High-Resolution Remote Sensing Images and the Effect of Sensitive Features on Classification Accuracy. *Sensors* 18, 7 (2018), 2013. [doi:10.3390/s18072013].