



Universidad de Jaén

Centro de Estudios de
Postgrado

IDENTIFICACIÓN DE BARRIOS DE INFRAVIVIENDAS (*SLUMS*) EN ARGELIA MEDIANTE TELEDETECCIÓN

Alumno/a: Rahmaní Abdelaziz Haizer

Tutores: Prof. D. Tomás Manuel Fernández Del Castillo

Prof. D. Francisco Javier Cardenal Escarcena

Dpto: Ingeniería Cartográfica, Geodésica y Fotogrametría

DICIEMBRE 2022



Universidad de Jaén

Centro de Estudios de
Postgrado

TRABAJO FIN DE MÁSTER

Máster Universitario en Ingeniería Geomática y Geoinformación

Estudiante:

Tutor/es

FERNANDEZ DEL
CASTILLO
TOMAS -
24234514N

Firmado
digitalmente por
FERNANDEZ DEL
CASTILLO TOMAS
- 24234514N
Fecha: 2022.12.12
20:58:34 +01'00'

CARDENAL
ESCARCENA
FRANCISCO
JAVIER -
08818663A

Firmado
digitalmente por
CARDENAL
ESCARCENA
FRANCISCO JAVIER
- 08818663A
Fecha: 2022.12.12
20:49:39 +01'00'

Fdo. RAHMANI ABDELAZIZ HAIZER

Fdo. TOMÁS M. FERNÁNDEZ DEL CASTILLO /

Fº JAVIER CARDENAL ESCARCENA

Jaén, 12 de diciembre de 2022



Tablas	5
1 RESUMEN/ABSTRACT	6
2 INTRODUCCIÓN	8
3 OBJETIVOS	9
4 ANTECEDENTES	10
4.1 El problema de los barrios de infravivienda en el mundo y en Argelia.....	10
4.2 Aplicación de la teledetección en la identificación de barrios de infravivienda	11
5 MÉTODOS.....	13
5.1 Zonas de estudio	13
5.1.1 La zona de Es Senia.....	13
5.1.2 La zona de Orán.....	16
5.1.3 La zona de Argel	17
5.2 Materiales.....	19
5.2.1 Imágenes	19
5.2.2 Software y hardware	20
5.3 Metodología	22
5.3.1 Pre-procesamiento de los datos.....	22
5.3.2 Clasificación	33
5.3.3 Aplicación de filtros texturales.....	40
5.3.4 Estimación de la probabilidad de barrios de infravivienda (<i>slums</i>)	50
5.3.5 Validación de los resultados.....	52
6 Resultados y discusión.....	58
6.1 Mapas de probabilidad a los slums	58
6.1.1 Zona piloto (Es Senia)	58
6.1.2 Zona de Orán	59
6.1.3 Zona de Argel.....	60
6.2 Validación	62
6.2.1 Zona piloto (Es Senia)	62
6.2.2 Zona de Orán	63
6.2.3 Zona de Argel.....	64
7 Conclusiones.....	66
8 Bibliografía.....	68



Tabla de figuras

Figura 1: Mapa de la zona de Es-Senia.	13
Figura 2: Mapa de la región de Orán	16
Figura 3: Mapa de la región de Argel.....	17
Figura 4. Diagrama de flujo del pre-procesamiento y clasificación de las imágenes de Sentinel-2.....	22
Figura 5. Ventana de Subset	23
Figura 6: Ventana de Resampling.	24
Figura 7: Imagen de la zona de Orán-Es Senia en 2022	25
Figura 8: Imagen de la zona de Orán-Es Senia de 2015.	25
Figura 9: Diagrama de flujo del pre-procesamiento de las imágenes de Sentinel-1	26
Figura 10. Ventana de TOPSAR Split	26
Figura 11: Ventana de Apply Orbit File	27
Figura 12: Ventana de Create Stack.....	28
Figura 13: Ventana de Back Geocoding	28
Figura 14: Ventana de Enhanced Spectral Diversity.....	29
Figura 15: Ventana de Coherence Estimation.....	30
Figura 16: Ventana de TOPS Deburst.....	30
Figura 17: Ventana de Multilooking.....	31
Figura 18: Ventana de Terrain Correction.....	32
Figura 19. Imagen de coherencia entre las dos imágenes de radar de Sentinel-1.....	32
Figura 20: Definición de los polígonos o vectores de entrenamiento.....	34
Figura 21: Ventana de clasificación random forest (entrada de datos).....	35
Figura 22: Ventana de clasificación random forest (parámetros).....	35
Figura 23: Clasificación general de las coberturas del suelo mediante random forest en la zona de Es Senia.	36
Figura 24: Ventana de Band maths, donde se aplica la reclasificación	37
Figura 25: Máscara de la zona urbana según Sentinel-2 (1: Urbano; 0; No urbano).....	37
Figura 26: Ventana de Band Maths, donde se aplica la máscara con una operación de multiplicación	38
Figura 27: Banda 8 de la imagen Sentinel-2 enmascarada en las zonas urbanas en 2022.....	38
Figura 28: Imagen de Sentinel-1 clasificada con random forest	39
Figura 29: Máscara de la zona urbana según Sentinel-1 (1: Urbano; 0; No urbano).....	39
Figura 30: Diagrama de flujo del análisis textural y morfología matemática.....	40
Figura 31: Ventana de los Filtros morfológicos.....	41
Figura 32: Resultado de aplicar el filtro de dilatación sobre la máscara y la imagen en la zona de Es Senia.	42
Figura 33: Resultado de aplicar la erosión en la imagen y la máscara en la zona de Es Senia.	42
Figura 34: Resultado de aplicar la apertura en la imagen y la máscara en la zona de Es Senia.....	43
Figura 35: Resultado de aplicar el cierre en la imagen y la máscara en la zona de Es Senia.	44
Figura 36: Ventana de filtros texturales GLCM	46
Figura 37: Resultados de la aplicación de filtros GLCM con matriz de 11x11 sobre la banda 8 en la zona piloto de Es Senia.	47
Figura 38: Resultados de la aplicación de filtros GLCM con matriz de 15x15 sobre la banda 8 en la zona piloto de Es Senia.	47
Figura 39: Resultados de la aplicación de filtros GLCM con matriz de 21x21 sobre la banda 8 en la zona piloto de Es Senia.	48
Figura 40: Resultados de la aplicación de filtros GLCM sobre imágenes con distinto pre-procesado en la zona piloto.	48



Figura 41: Resultados de la aplicación de filtros GLCM sobre imagen de la entropía en la apertura	49
Figura 42: Resultados finales de la aplicación de filtro GLCM de disimilitud y post-proceso (apertura) en la zona piloto.	49
Figura 43: Resultados finales de la aplicación de filtro GLCM de varianza y post-proceso (apertura) en la zona piloto	49
Figura 44: Imagen preliminar de la probabilidad de slums.....	50
Figura 45: Imagen final de la probabilidad de slum.....	51
Figura 46: Mapa de probabilidad de slums en la zona piloto de Es senia.....	51
Figura 47: Polígonos digitalizados sobre la imagen de la banda 8.	52
Figura 48: Ventana de la herramienta de Puntos aleatorios dentro de polígonos	53
Figura 49: Capas de puntos aleatorio obtenidos para la validación en la zona piloto de Es Senia. En amarillo los puntos obtenidos como slum y en rojo como no slum.	53
Figura 50: Ventana de Point Sampling Tool	54
Figura 51: Tablas con los puntos de validación, el valor de probabilidad y la verdad terreno.....	54
Figura 52: Script de R para la obtención de la matriz de confusión y la curva ROC	57
Figura 53: Curva ROC y valor del área bajo la curva (AUC) en la zona piloto	57
Figura 54: Filtros texturales en la zona piloto mediante distintos operadores e imagen en color real (abajo)	58
Figura 55: Mapa de peligrosidad de slums en la zona piloto de Es Senia	59
Figura 56: Filtros texturales en la zona de Orán mediante distintos operadores	60
Figura 57: Mapa de peligrosidad de slums en la zona completa de Oran.....	60
Figura 58: Filtros texturales en la zona de Argel mediante distintos operadores.....	61
Figura 59: mapa de peligrosidad de slums en la zona de Argel	61
Figura 60: Puntos de muestreo para validación en la zona piloto	62
Figura 61: Curva ROC y resultado de la AUC en la zona piloto.....	62
Figura 62: Puntos de muestreo para validación en la zona de Oran.....	63
Figura 63: Curva ROC y resultado de la AUC en la zona de Orán	64
Figura 64: Puntos de muestreo para validación en la zona de Argel	64
Figura 65: Curva ROC y resultado de la AUC en la zona Argel	65

Tablas

Tabla 1: Distribución de la población activa	15
Tabla 2: Sector de la actividad económica.....	15
Tabla 3: Rama de la actividad económica.....	15
Tabla 4: Resolución de las bandas de las imágenes de Sentinel-2.	24
Tabla 5: Curva ROC y resultado de la AUC en la zona piloto	63



1 RESUMEN/ABSTRACT

Los barrios de infravivienda (*slums*) son un problema a nivel mundial y en particular en los países en desarrollo como Argelia. Por lo tanto, su identificación mediante diversas técnicas como la teledetección, es importante como primer paso para la actuación por parte de las autoridades. En este trabajo se propone en primer lugar la clasificación de imágenes de Sentinel-1 y 2 para la identificación de *slums* en Argelia. La clasificación supervisada con bosques aleatorios permite detectar las zonas urbanas sobre todo en imágenes de Sentinel-2, pero no entre distintos tipos de zonas urbanas como los *slums*. Por ello, se opta por hacer máscaras de las zonas urbanas y aplicar sobre ellas técnicas de análisis textural como los filtros de co-ocurrencia (GLCM) y morfología matemática. Los mejores resultados se obtienen con los filtros de entropía y, especialmente, de disimilitud, apoyados con operaciones de apertura y cierre. A partir de las imágenes filtradas, se realizan mapas de probabilidad de *slums* en una zona piloto en Es Senia, así como en Orán y Argel, la capital del país. Los mapas se validan mediante la matriz de confusión y el área bajo la curva ROC, proporcionando una fiabilidad del 80-90%.

Palabras clave: barrios de infravivienda (*slums*); cartografía; teledetección; análisis textural; Argelia.



Abstract

Slums are a problem worldwide and particularly in developing countries such as Algeria. Therefore, its identification through various techniques such as remote sensing is important as the first step for action by the authorities. In this work, the classification of Sentinel-1 and 2 images using random forest algorithm for the identification of slums in Algeria is proposed. The classification allows the detection of urban areas especially in Sentinel-2 images, but not between different types of urban areas such as slums. Therefore, it is decided to make masks of urban areas and apply textural analysis techniques such as co-occurrence filters (GLCM) and mathematical morphology. The best results are obtained with entropy filter and, especially, with dissimilarity filter, supported by opening and closing operations. Based on the filtered images, slum probability maps are made in a pilot area in Es Senia, as well as in Oran and Algiers, the country's capital. The maps are validated using the confusion matrix and the area under the ROC, providing a reliability of 80-90%.

Keywords: slums; mapping; remote sensing; textural analysis; Algeria.



2 INTRODUCCIÓN

En Argelia, el rápido ritmo de urbanización ha dejado a las ciudades incapaces de satisfacer la demanda de viviendas y servicios asequibles. Como resultado, las zonas deprimidas espontáneamente en las que se desarrollan barrios de infravivienda ("slums") se han extendido rápida y desordenadamente y, en su mayoría, de forma anárquica.

El problema de la vivienda precaria en Argelia no es nuevo. Se remonta a la época colonial, cuando aparecieron los primeros slums en 1930. En la actualidad, estas aglomeraciones siguen imponiéndose en el paisaje urbano al desarrollarse en la periferia de las grandes ciudades.

Más recientemente, entre 1990 y 2000, muchas familias huyeron de las zonas rurales aisladas y precarias por razones de seguridad y se instalaron cerca de las grandes ciudades, como Argel, Orán o Annaba, donde en (2003) se contabilizaron 24.000 viviendas precarias de un total de 544.000.

Sin embargo, las estimaciones de población de estas zonas siguen siendo difíciles porque las comunidades desfavorecidas no están incluidas en los mapas o datos oficiales. Las organizaciones nacionales e internacionales, las autoridades y las ONG necesitan urgentemente datos geoespaciales con la localización de estos barrios, para evaluar la situación y realizar esfuerzos específicos para mejorar las condiciones de vida. Por otra parte, los habitantes de los barrios marginales también quieren figurar en estos datos para ser realmente reconocidos como ciudadanos y tener voz.

La reabsorción de los slums se ha convertido en un proceso importante en Argelia, y se está convirtiendo en un tema de interés para varias disciplinas, planteándose como objetivo inicial la evaluación de este fenómeno y el seguimiento espacio-temporal de las ciudades. Entonces, la pregunta que surge es ¿cómo detectar la expansión de un poblado chabolista? Evidentemente hay muchas respuestas, es decir muchas formas de estudiarlo, pero una de ellas es la teledetección. La teledetección es una propuesta interesante en los tiempos actuales ya que proporciona imágenes de satélite continuas con una resolución adecuada para caracterizar los territorios y su evolución. Dentro de sus múltiples aplicaciones, la teledetección, es decir, la captura y el análisis de las imágenes de satélite permite destacar los cambios y analizar la dinámica urbana y periurbana (Ding et al., 2007; Hu y Lo, 2007; Chi et al., 2007; Maseket al., 2000), y en concreto para identificar y monitorizar los slums.

La teledetección se define como "el conjunto de conocimientos y técnicas utilizados para la observación, el análisis, la interpretación y la gestión del medio ambiente a partir de mediciones e imágenes obtenidas mediante plataformas aéreas, espaciales, terrestres o marinas".



En este trabajo, nos basamos en imágenes de satélite del programa Copernicus, en concreto de Sentinel-1 y Sentinel-2, que abarcan escenas de la región de Orán y de Argel. La interpretación de las imágenes se basa también en los resultados del tratamiento digital (clasificación supervisada, análisis de texturas GLCM, filtros morfológicos, etc.) con el fin de cartografiar los barrios de chabolas y sus patrones de crecimiento.

3 OBJETIVOS

El objetivo principal de este trabajo es la identificación de barrios de infravivienda (*slums*) en Argelia, mediante técnicas de teledetección espacial.

Para alcanzar este objetivo principal, se proponen algunos objetivos específicos:

- Ensayar técnicas de análisis textural y morfología matemática para cartografiar la probabilidad de que un barrio o asentamiento humano sea un *slum*.
- Evaluar su validez en casos concretos de ciudades de Argelia, con vistas a un conocimiento preliminar de la situación y a la promoción de medidas para su mejoramiento.



4 ANTECEDENTES

4.1 El problema de los barrios de infravivienda en el mundo y en Argelia

El problema de los slums o de la precariedad de la vivienda es uno de los mayores problemas del siglo XXI, sin embargo, a pesar del paso de los años y de las soluciones encontradas, salvo que sigue siendo un problema muy frecuente en los países en desarrollo y entre los países más afectados de África y de la cuenca mediterránea está Argelia.

Este problema se remonta a la época colonial, habiendo aparecido ésta en 1930, y sigue imponiéndose hoy en el paisaje urbano al desarrollarse en la periferia de las ciudades, normalmente asociado a la emigración de las zonas rurales a las ciudades en busca de oportunidades y seguridad.

En un documento dedicado a la reestructuración de las viviendas precarias, el Ministerio de Vivienda precisa que durante la década negra que vivió Argelia entre 1990 y 2000, por razones de seguridad, muchas familias huyeron de las zonas rurales aisladas y precarias para instalarse cerca de las grandes ciudades, como Argel, Orán o Annaba.

Esto llevó al registro (en 2003) de 24.000 viviendas precarias de las 544.000 existentes en el territorio nacional. En 2010, la cifra para todo el país era de 554.000, incluyendo todos los tipos de infravivienda, con un 60% de estas infraviviendas situadas en zonas urbanas y un 40% en zonas rurales (El-Moudjahid PressCentre, 2010).

Sobre la base de estos datos, se puede considerar que los barrios de infravivienda, que incluyen slums y emplazamientos espontáneos no equipados, se concentran en las zonas urbanas y rurales, así como en las grandes ciudades, y cubren, según el mismo estudio (Banco Mundial, Ministerio de Vivienda), una superficie total de 17.000 ha en el territorio argelino. Dada su elevada tasa de crecimiento demográfico (más del 4,5% anual), Argel, Tipaza, Blida, Constantina, Orán y Annaba son las más afectadas por la precariedad de la vivienda.

Las encuestas realizadas por el Ministerio de Vivienda en el marco del proyecto de lucha contra la precariedad de la vivienda revelan que el 40% de los municipios están afectados, es decir, 270 edificios insalubres por municipio en la periferia urbana y el centro de las ciudades, o en terrenos no edificables (en la ladera de una montaña, cerca de un barranco o en una zona industrial).

Cabe señalar que el 74% de estas construcciones situadas en el centro y el oeste del país son construcciones básicas (chozas y chapas sin estructura de soporte), frente al 26% de estructuras similares situadas en el este.



A partir de esta información, se puede afirmar que aproximadamente el 40% de los edificios existentes se construyeron antes de 1962, periodo que corresponde a los campos de reagrupación y reasentamiento. A pesar de esta alarmante situación, que corresponde a 554.000 favelas en todo el país, y en comparación con los países árabes, Argelia tiene una de las tasas más bajas de viviendas precarias, que no supera el 10% del parque nacional de viviendas.

En comparación con los países árabes, Argelia tiene una de las tasas más bajas de viviendas precarias, que no superan el 10% del parque nacional de viviendas. Así, en cifras, "este tipo de vivienda es el 39% en Egipto, el 32% en Marruecos y Libia y el 94% en Mauritania" (Centro de Prensa El-Moudjahid, 2010).

4.2 Aplicación de la teledetección en la identificación de barrios de infravivienda

El desarrollo de la vivienda informal es una preocupación creciente para los servicios del Estado y las comunidades, que enfrentan varios desafíos: seguridad de bienes y personas (ocupación de áreas con riesgos ambientales, sobrepoblación, insalubridad, delincuencia, traficantes de sueño urbanístico...), lucha contra la pobreza, ilegalidad inmigración, etc. Pero es un fenómeno difícil de medir porque evoluciona de forma rápida y discreta. Además, enviar agentes al campo para realizar encuestas suele ser problemático por múltiples razones.

Por lo tanto, las imágenes de satélite parecen ser una oportunidad real para comprender mejor la extensión de las áreas en cuestión, detectar nuevas construcciones y comprender la tasa de cambio. Así, en los últimos años se ha desarrollado algunos trabajos que aplican las técnicas de teledetección al estudio de los barrios de infravivienda, *slums* (Kuffer et al., 2016). En algunas publicaciones ha habido varias iniciativas para el mapeo y detección de barrios marginales, y entre los métodos más utilizados se encuentran el método de cuadrícula aplicado en la Guayana Francesa (Referencia, año):

Inicialmente, las zonas prioritarias se definen con las comunidades involucradas, en forma de cuadrados de 200 metros de lado, lo que evita la dependencia de la zonificación administrativa y permite enfocar el análisis en los sectores objetivo. Para cada territorio objetivo, se realizó un mapeo del edificio, utilizando los repositorios disponibles en la Guayana Francesa (vector PCI y capa de edificio del BD Topo) con el fin de proporcionar una matriz de referencia, superpuesta a la ortofotografía más detallada disponible. En el sector de Cayena se utilizó una ortofoto de 20 cm de resolución de 2015.

Otro método fue aplicado por Kohli (2015), en las ciudades de Nairobi y Ahmadabad. En este trabajo se analiza la aplicabilidad de la ontología genérica de barrios marginales de GSO y se explora la transferibilidad de un método de análisis orientado a objetos OOA para la clasificación de barrios marginales desarrollado en un área de imagen única a otras áreas de imagen. Para incluir el



conocimiento contextual, el conocimiento experto se codifica en el marco ontológico para clasificar los barrios marginales en los tres subconjuntos de la imagen GeoEye de la ciudad de Ahmedabad, India. Así, se identifican un conjunto de parámetros basados en imágenes como los más estables o transferibles para diferenciar los barrios marginales de las áreas que no son barrios marginales en todos los subconjuntos. El método se aplica a una imagen Quickbird de la ciudad de Pune, India. La integración del conocimiento experto sobre barrios marginales se realiza en forma de una ontología local de barrios marginales que también demuestra su aplicabilidad en un contexto diferente. Finalmente, se demuestra un método para cuantificar las incertidumbres de las interpretaciones expertas de los barrios marginales. Se examinan las delimitaciones realizadas en tres contextos diferentes y se estudian los factores responsables de las incertidumbres en la distinción de barrios marginales de zonas no marginales.

En el artículo que se toma de referencia en este trabajo (Wurm et al., 2017), se utiliza la textura de la imagen y los perfiles morfológicos para mapear barrios marginales en imágenes Sentinel-2A. Los diferentes tamaños de los respectivos descriptores espaciales (GLCM, perfiles morfológicos diferenciales) se prueban para la clasificación utilizando un clasificador de bosque aleatorio (random forest). Los resultados se interpretan en función de una evaluación de precisión basada en píxeles y en parcelas. Los mejores resultados de clasificación se han alcanzado en el nivel basado en píxeles con un valor de kappa de 81,65 para el conjunto de características combinadas con GLCM y perfiles morfológicos. A nivel de parcelas, los análisis muestran que se alcanzan precisiones más altas con tamaños de kernel grandes de los filtros (óptimo con tamaños de 17x 17) y que la detección es mejor para áreas grandes de slums. En trabajos posteriores de ese equipo, se han empleado redes neuronales convolucionales para la segmentación semántica de las imágenes que permite identificar los slums (Wurm et al., 2019).

5 MÉTODOS

5.1 Zonas de estudio

5.1.1 La zona de Es Senia

ES SENIA, antiguamente conocido como La Sénia durante la colonización francesa, es una daïra (distrito) y comuna (municipio) de la *wilaya* (provincia) de Orán en Argelia, situado en la periferia sur de Orán, con una superficie de 48,51 km² y una altitud de 77 m.



Figura 1: Mapa de la zona de Es-Senia.

Teniendo en cuenta un análisis realizado por la daïra de Es Senia (wilaya de Orán) con vistas a una posible actualización de un proyecto de reducción de la vivienda precaria en el douar (agrupación rural) árabe y el douar marroquí de la comuna de Ain el Beida. Tomando como modelo las otras provincias que tienen el mismo problema y que ya han tenido la experiencia de proyectos similares.

El entorno y el acceso a la población de Ain El Beida, que incluye los dos douars, está situada al oeste del municipio de Es Senia y al suroeste de la periferia de la ciudad de Orán, de la que dista ocho kilómetros. Con una superficie de 28,5 ha, los dos douars conocidos como "Douar Arabe y Douar Marocain" constituyen el núcleo urbano original alrededor del cual se han desarrollado una serie de programas de vivienda planificada de diferentes tipologías (vivienda agrícola rural, autoconstrucción, urbanización, semicolectiva, colectiva, social...).



Aunque geográficamente ocupa una posición central en la aglomeración de Ain El Beida, el lugar está cerrado debido a la presencia de dos grandes barrios de infravivienda que constituyen una ruptura física entre las dos entidades urbanas, por un lado, y entre los dos douars y la aglomeración, por otro. En cuanto a la topografía del lugar, los dos sitios tienen una pendiente que varía del 3 al 10%. Esta morfología irregular del lugar es la causa de las inundaciones durante los períodos de lluvia en la parte sur de la zona de estudio.

Además, algunos de los callejones han sido ocupados, lo que dificulta el tránsito de los peatones, ya que la gente tiende a instalarse donde puede, tratando de utilizar el mayor espacio posible. Hay unas 852 parcelas de diferentes formas y tamaños que van de 25 a 640 m², en las que el tipo de vivienda predominante es el haouch de una a dos habitaciones, a veces más (una vivienda formada por unas cuantas habitaciones dispuestas alrededor de un patio interior, a menudo en forma de U). El sistema de carreteras dentro del lugar es casi inexistente, y a menudo consiste en senderos y carriles estrechos y sinuosos, lo que da al distrito un aspecto enmarañado y anárquico.

Respecto a nuestra zona piloto, que es la zona de Essenia, presentaremos información cuantitativa para una mejor descripción de los lugares (Fuente). El análisis de la propia población incide en los aspectos demográficos, su perfil socioeconómico, los ingresos de los jefes de hogar, su grado de solvencia, etc. Este asentamiento tiene un total de 1061 hogares. La población total contabilizada en el área de estudio fue de 6.509 personas, divididas en 1.046 familias, de las cuales el 91,14% eran familias nucleares o simples y el 7,45% eran familias polinucleares o compuestas. Una familia nuclear se define como una familia compuesta por una pareja y sus hijos solteros, y una familia polinuclear se define como una familia compuesta por dos o más parejas y sus hijos solteros. Observamos que la familia nuclear es la más dominante, lo que explica que la población del sitio estudiado sea joven. Hay 79 familias polinucleares que representan el 7,45% de todas las familias. La reducción del tamaño de las familias observada en los dos precarios douares refleja, por un lado, la tendencia argelina, pero también la impone este tipo de vivienda, cuya superficie el promedio es generalmente muy bajo (25 m²) por familia; y por otra parte, la inestabilidad de esta población, que es en gran parte resultado del éxodo.

Por su parte, la población activa (población ocupada y desempleada) asciende a 1.047 personas, el 16,08% de la población total, lo que da una tasa de dependencia de 6,2. La población ocupada asciende a 688 personas, lo que supone una tasa de ocupación del 65,71%. Los parados o parados son alrededor de 359 personas, lo que da una tasa de desempleo de 34.29%, como se muestra en la siguiente tabla 1:



Población activa	Numero	%
Empleado	688	65,71
STR	359	34,29

Tabla 1: Distribución de la población activa

En la tabla 2 se presenta la distribución por sectores, donde se observa como el sector terciario es el mayoritario.

Sector de actividad económica.	Fuerza laboral	%
Sector primario	14	2,03
Sector secundario	121	17,59
Sector terciario	553	80,38
Total	688	100

Tabla 2: Sector de la actividad económica

La distribución de los empleados en cada área de la actividad económica se presenta en la tabla 3.

Rama de actividad económica.	Fuerza laboral	%	Total %
Agricultura y ganadería	14	2,03	2.03
Industria	20	2,91	17.59
Construcción y obras públicas	101	14,68	
Administración	142	20,64	80.38
Transporte	43	6,25	
Educación	8	1,16	
Salud	5	0,73	
Tiendas y servicios	355	51,60	
Total	688	100	100

Tabla 3: Rama de la actividad económica

5.1.2 La zona de Orán

La wilaya de Orán, es la segunda ciudad más grande de Argelia y una de las ciudades más importantes del Magreb, tiene una superficie de 64 km². Es una ciudad portuaria sobre el Mar Mediterráneo, ubicada en el noroeste de Argelia, a 432 km de la capital Argel, sus coordenadas 35° 42' 10" norte, 0° 38' 57" oeste con una altitud (Min. 0m, Max. 580 m). Orán se compone administrativamente de 26 municipios y 09 daïras subdivididas. En 2014, la población residente se estima en más de 1.822.000 habitantes.



Figura 2: Mapa de la región de Orán

La observación de la estructura por edades de la población residente en el territorio de la wilaya en 2014 pone de manifiesto la gran proporción de habitantes jóvenes que alberga con un 45,56% de la población menor de 24 años. Una gran proporción de personas en edad de trabajar (46,35%), así como la proporción de personas mayores sigue siendo baja (8,08%).

La wilaya de Orán registra una tasa media de crecimiento anual del 9% en términos de empleo, que sigue siendo sensiblemente superior a la de la población. Este crecimiento ha permitido reducir considerablemente la tasa de desempleo, que es del 10,89%, aunque la tasa media de desempleo de la wilaya, que es del 10,89%, parece satisfactoria, sigue siendo matizada a nivel de los municipios: sólo los municipios de Es Seina, Orán y Bir El Djir presentan un resultado inferior al de la wilaya.

Los demás municipios presentan tasas superiores a la media de la wilaya, con diferencias de hasta 5 puntos, y son los municipios del sur de la wilaya donde el desempleo sigue siendo elevado. Contrariamente a lo que presentan los poderes públicos, las soluciones aportadas siguen siendo simples soluciones de espera y el tratamiento del desempleo no es económico sino fundamentalmente social.



En cuanto a la urbanización, la wilaya de Orán es una de las zonas más urbanizadas del país, con una tasa de urbanización superior al 90%. Cabe señalar que las reformas introducidas en el sector de la vivienda han acelerado la producción del entorno construido, lo que permite una amplia gama de intervenciones y una diversidad de promotores que sin duda se reflejarán en los resultados cualitativos y cuantitativos.

La ruralización que sigue golpeando duramente a los grandes centros urbanos de la wilaya de Orán, encabezados por la muy comercial daïra de Es Senia, situada en el suroeste de la ciudad de Orán. Una encuesta, que informa de 20.000 chabolas repartidas por las cuatro esquinas que rodean la ciudad. Dicha cifra representa el 50% de las chabolas situadas en los 26 municipios y nueve daïras que componen el territorio de Orán.

Peor aún, la mayoría de estas "favelas" al estilo brasileño están situadas en zonas que no cumplen ninguna norma de seguridad. El mayor número de construcciones ilegales se ha detectado en los municipios de Sidi Chahmi, Es-Sénia y Haï Bouâmama, hassi bounif, bir el djir, es sebkha, sidi el bachir, cheklaoua, bab el hamra, regencia de hayet, etc.

5.1.3 La zona de Argel

La wilaya de Argel es una subdivisión administrativa de Argelia ubicada en Argel. Corresponde a la ciudad de Argel e incluye el centro de esta última así como sus distintos distritos. La wilaya se divide en 57 municipios o comunas, cuya capital (prefectura) es Argel, agrupados en 13 distritos administrativos resultantes de la división de 2 de agosto de 1997; en esta se integraron distritos anteriormente pertenecientes a las wilayas limítrofes de Tipaza, Blida y Boumerdès, caracterizadas por su muy baja densidad de poblamiento en comparación con las de la wilaya de Argel.



Figura 3: Mapa de la región de Argel



La wilaya de Argel es la más poblada de Argelia con 3.154.792 habitantes, aunque también es la más pequeña, con una superficie de 1.190 km². Supone el 10% de los 44 millones de argelinos. Este peso demográfico relativamente modesto, comparado con el de las grandes aglomeraciones del sur del Mediterráneo como El Cairo, Túnez, Casablanca, no expresa el considerable peso económico, administrativo y cultural de la ciudad a escala nacional. Argel alberga el aeropuerto más grande del país, con una capacidad de 12 millones de pasajeros/año, pero también el puerto más grande del país, que representa el 47% de toda la actividad portuaria del país. El puerto de Argel manipula un volumen de 500.000 contenedores al año. La ciudad de Argel es también y sobre todo el centro de empleo administrativo más importante del país, con el 23% del total de la plantilla administrativa. Es sede de las dos mayores universidades del país, con una población estudiantil superior a los 200.000 individuos, pero también es la ciudad donde se concentra la mayor cantidad de centros de investigación y laboratorios científicos del país. Finalmente, alberga las zonas industriales y de actividad económica más importantes a nivel nacional: cuatro zonas industriales repartidas en 1.378 ha y 26 zonas de actividad repartidas en 257,29 ha. Estas funciones de metrópoli nacional juegan un cierto papel en la concentración de puestos de trabajo vinculados a los roles de mando y poder político y económico en la capital, sin explicar, sin embargo, la distribución espacial de los puestos de trabajo en la aglomeración.

En cuanto al paisaje urbano de la wilaya de argel nos interesa más la problemática de los barrios marginales que rodean la ciudad de argel un poco más en todos los demás rincones de la wilaya. Nótese que la wilaya de argel ha puesto en marcha el programa de capital con cero favelas, pero nos vamos a interesar un poco en el destino de las favelas más grandes, entre ellos el famoso "Hai Eremli", el primer barrio marginal erradicado en Argel.

Extendiéndose sobre un área de 54 hectáreas en Gué de Constantine en Argel, "Hai Eremli" tenía fama de ser el barrio marginal más grande de la capital e incluso del país, que apareció en 1994. Este sitio llegó a estar ocupado por 4.000 familias. "Hai El Eramli" fue erradicado en 2015, dando paso a proyectos estructurales en la capital.

En mayo de 2016 se erradicaron otros dos barrios marginales, a saber, "Oued El Hamiz" en Rouiba (1800 familias) y "Hai Benboulaid" en Hussein-Dey (400 familias). En cuanto al barrio marginal de "Hai Benboulaid" en El Magharia en la Daïra de Hussein-Dey, fue eliminado en presencia del director regional de la organización de las Naciones Unidas para el Hábitat (ONU-hábitat), Alioune Badiane, quien saludó los esfuerzos de Argelia. en el campo de la vivienda, particularmente en términos de eliminación de barrios marginales.



Por su parte, los ocupantes de "Hai El Kerouch" en Reghaia (1500 familias), de "Kariet Echouk" en Gué de Constantine (1200 familias), de la mezquita "Hai Dergana" en Bordj El Kiffan (400 familias), de "Hai El Hoffra" en Oued Esmar (1200 familias), "du Hai Coco Plage" en Bordj El Bahri (186 familias), "Hai Sidi Messaoud" ubicado cerca de la avenida Mohamed Belouizdad (400 familias), "Hai Omar Ouahib" en Bologhine (400 familias) y "Hai Edoum" en Bordj El Kiffan (150 familias), todas han sido realojadas en hogares dignos.

Finalmente, otros barrios marginales, a saber, "Hai Ain Zeboudja" en Bouzaréah (226 familias), "Hai Boualem Rouchai" en Ouled Fayet (2014 familias) y "Hai Grand Rocher" en Ain Bénian (165 familias) fueron arrasados simultáneamente el 26 de diciembre de 2017 cediendo 13 hectáreas de terreno en beneficio de varios proyectos estructurantes en los sectores de hidráulica y equipamiento público.

5.2 Materiales

5.2.1 Imágenes

El programa Copérnico es un programa de observación de la Tierra desarrollado por la Agencia Espacial Europea (ESA) y centrado en diferentes aspectos de la observación de la Tierra, mediciones atmosféricas y oceánicas, se compone de 6 misiones y constelaciones (<https://www.copernicus.eu>), los Sentinel, de los que en este trabajo se van a utilizar las imágenes de Sentinel-1 y Sentinel-2.

Copernicus Sentinel-1: Es un radar de apertura sintética (SAR) de imágenes diurno y nocturno para servicios terrestres y oceánicos. Los satélites gemelos Sentinel-1A y Sentinel-1B fueron lanzados el 3 de abril de 2014 y el 25 de abril de 2016, respectivamente.

Para este estudio se han descargado desde la plataforma de Copernicus Open Access Hub (<https://scihub.copernicus.eu/>) dos imágenes de Sentinel 1 A y 1B de Single-Look Complex (SLC) e Interferometric Wide Swath Mode (IW) para las fechas 27-05-2022 y 21-11-2019 y la zona de Orán-Es Senia.

- S1A_IW_SLC__1SDV_20220527T180153_20220527T180223_043400_052EC7
- S1B_IW_SLC__1SDV_20191121T180052_2022_019029_023E93_6DDD

Copernicus Sentinel-2 es un satélite óptico de alta resolución para servicios en tierra. Proporciona imágenes de la vegetación, la cubierta terrestre y acuática, las vías navegables interiores y las zonas costeras, por ejemplo. Los satélites gemelos Sentinel-2A y Sentinel-2B fueron lanzados el 22 de junio de 2015 y el 7 de marzo de 2017, respectivamente.



Para este estudio se han descargado imágenes de Sentinel-2 de nivel 2A (corregidas atmosféricamente) para las fechas 27-05-2022 y 26-07-2022 de la zona de Orán-Es Senia

- S2A_MSIL2A_20220527T103631_N0400_R008_T30SYE_20220527T183616
- S2A_MSIL1C_20150726T105026_N0204_R051_T30SYE_20150726T105856

En cuanto a la zona de Argel, las imágenes de Sentinel-2 descargadas son de las fechas 28-06-2022 y 05-04-2018 :

- 'S2B_MSIL2A_20220628T102609_N0400_R108_T31SEA_20220628T133556 2
- S2A_MSIL2A_20180405T103021_N0207_R108_T31SEA_20180405T141122

5.2.2 Software y hardware

La plataforma de aplicaciones Sentinel (SNAP) es una arquitectura común para todos los kits de herramientas Sentinel. El software ha sido desarrollado por Brockmann Consult, Skywatch, Sensar y C-S. La arquitectura SNAP es ideal para el procesamiento y el análisis de la Observación de la Tierra (EO) gracias a las siguientes innovaciones tecnológicas: escalabilidad, portabilidad, plataforma cliente rica y modular, abstracción genérica de datos EO, gestión de memoria en mosaico y marco de procesamiento de gráficos. Además, SNAP y los kits de herramientas individuales de Sentinel admiten muchos sensores que no son de Sentinel.

ESA/ESRIN proporciona la herramienta de usuario SNAP a la comunidad de observación de la Tierra de forma gratuita. Se ha publicado la última actualización de SNAP 9.0, una herramienta en línea de acceso gratuito que ayuda a aprovechar el enorme potencial de los datos proporcionados por los satélites de teledetección.

La plataforma permite acceder a las imágenes de Sentinel-1, Sentinel-2 y Sentinel-3 del programa Copérnico de la Unión Europea, así como a las misiones Earth Explorer de la ESA, como SMOS. También integra datos de muchos satélites operados por socios internacionales, conocidos como "misiones de terceros o terceras partes".

SNAP cuenta con una interfaz gráfica fácil de usar y también puede funcionar desde la línea de comandos. Por lo tanto, es accesible para personas con poca experiencia en codificación y programación, así como para expertos en análisis de datos.

Por su parte, QGIS es un Sistema de Información Geográfica (GIS) fácil de usar distribuido bajo la Licencia Pública General GNU. Es un proyecto oficial de la Open Source Geospatial Foundation (OSGeo), anteriormente conocida como Quantum GIS.



El software fue diseñado en 2002 y maduró alrededor de 2009-2010. Desde entonces, su uso ha sido exponencial y hoy podemos decir que QGIS combina el acceso gratuito con el rendimiento para convertirse en un referente en el mundo de los SIG. Por otro lado, QGIS cuenta con una gran comunidad que le otorga una estabilidad, un campo funcional y una escalabilidad equivalente o incluso superior a las soluciones de pago del mercado como ArcGis, MapInfo o GeoConcept.

El uso de QGIS es relevante en muchos oficios y temas: Urbanismo, Ordenamiento Territorial, Medio Ambiente, Fibra Óptica FTTH, Redes (agua, saneamiento), telecomunicaciones, transporte, energía, inmobiliario, salud, geomarketing, Diseño de oficinas, EPCI (autoridades locales), seguros, bancos, distribución masiva, mayoristas, industrias, asociaciones. QGIS ofrece un número cada vez mayor de posibilidades, a través de sus funciones internas y sus extensiones. Puede ver, administrar, editar, analizar datos y componer mapas para imprimir.

En cuanto al hardware, se trata de una computadora Lenovo I5 quinta generación. Aunque ha permitido realizar el trabajo, es claro que se requiere una computadora más potente para la realización de todas las operaciones en un menor lapso de tiempo.

5.3 Metodología

La metodología comprende en primer lugar el pre-procesamiento de los datos, seguido por la clasificación y obtención de máscaras; a continuación, se aborda la parte más importante que es la aplicación de filtros texturales para la determinación de la probabilidad de encontrar barrios de infravivienda, seguido de un post-procesamiento para mejorar la calidad visual de los resultados; finalmente se realiza una validación mediante la matriz de confusión y la curva AUC-ROC.

5.3.1 Pre-procesamiento de los datos

En este apartado se presenta el pre-procesamiento tanto de las imágenes de Sentinel-1 como de Sentinel-2.

5.3.1.1 Pre-procesamiento de imágenes de Sentinel-2

La preparación de las imágenes de Sentinel-2 para su posterior explotación es un proceso algo más sencillo que en las imágenes de Sentinel-1, que se explica en los apartados siguientes. Se resume en la figura 4 (en la que también se muestran las operaciones subsiguientes de clasificación y generación de la máscara urbana).

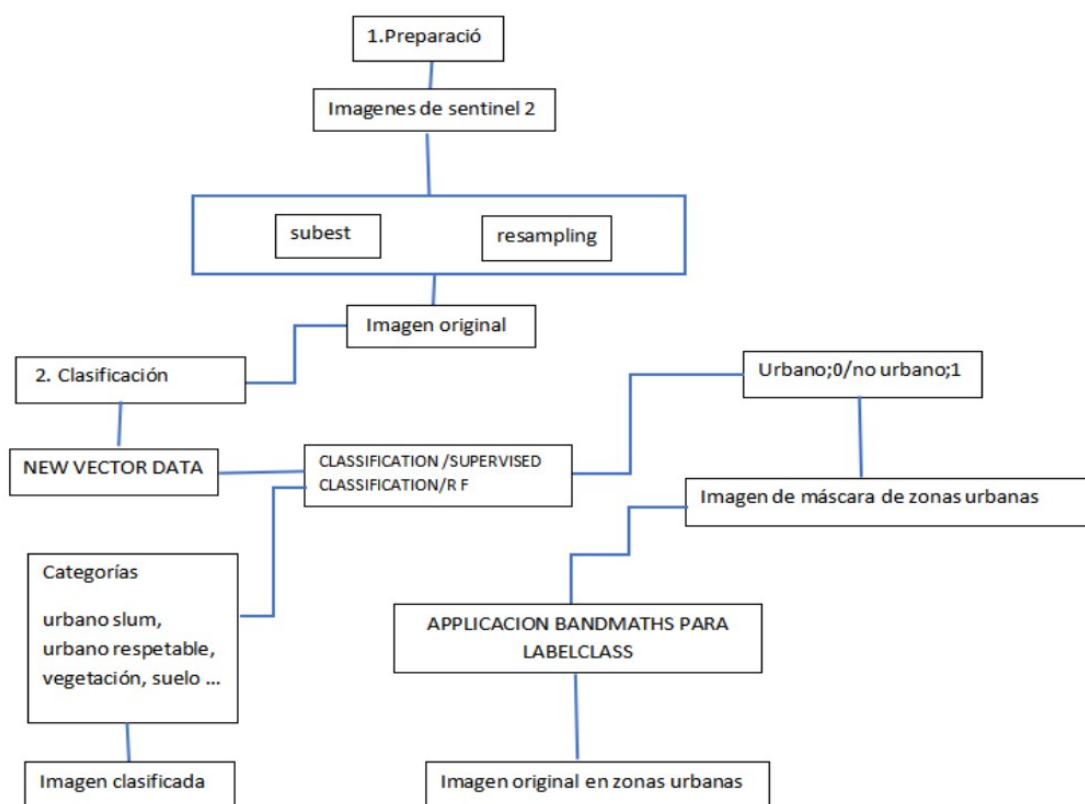


Figura 4. Diagrama de flujo del pre-procesamiento y clasificación de las imágenes de Sentinel-2

5.3.1.1.1 Subconjunto (subset)

Con esta herramienta (Figura 5) se crea un subconjunto o recorte del producto, ya que no se necesita todo el ámbito. Se puede definir un área según nuestras necesidades, o incluso saltarse el paso de Subconjunto y mantener toda la extensión.

Raster → Geometric → Subset.

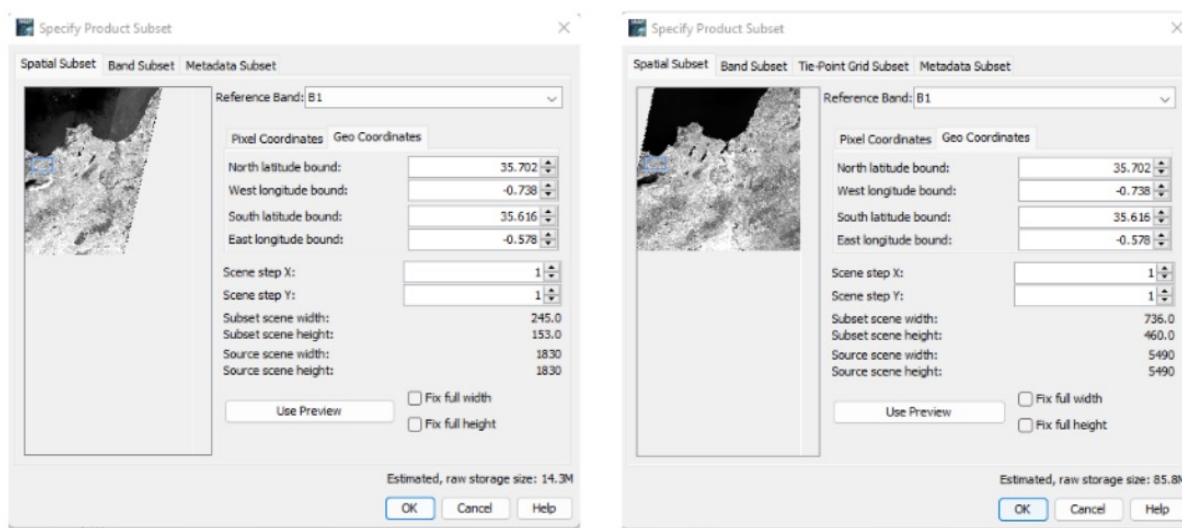


Figura 5. Ventana de Subset

5.3.1.1.2 Remuestreo (resampling)

El remuestreo permite presentar y procesar de manera uniforme datos de diferente resolución espacial. Así, los datos del sensor óptico de Sentinel-2, se proporcionan con tres tamaños de píxel diferentes: 10, 20 y 60 metros (ver Tabla 1). Como el software SNAP de la ESA requiere la misma resolución espacial para cada banda de la imagen de origen para su completa visualización y para muchas de funciones, se procede al remuestreo de la imagen.

La operación de remuestreo se realiza mediante la siguiente herramienta:

Raster → Geometric → Resampling

En la pestaña Remuestreo (Figura 6), en la opción "Establecer el tamaño del producto remuestreado", se selecciona la opción: Por banda de referencia del producto de origen, y allí marcar la B2. De este modo, volveremos a muestrear todas las bandas a la máxima resolución de 10 m.



Bandas	Longitud de onda (μm)	Resolución (m)
Banda 1 – Aerosol	0.43 – 0.45	60
Banda 2 – Blue	0.45 – 0.52	10
Banda 3 – Green	0.54 – 0.57	10
Banda 4 – Red	0.65 – 0.68	10
Banda 5 – Red Edge 1	0.69 – 0.71	20
Banda 6 – Red Edge 2	0.73 – 0.74	20
Banda 7 – Red Edge 3	0.77 – 0.79	20
Banda 8 – Near Infrared (NIR 1)	0.78 – 0.90	10
Banda 8A – Near Infrared (NIR) 2	0.85 – 0.87	20
Banda 9 – Water vapour	0.93 – 0.95	60
Banda 10 – Cirrus	1.36 – 1.39	60
Banda 11 – SWIR 1	1.56 – 1.65	20
Banda 12 – SWIR 2	2.10 – 2.28	20

Tabla 4: Resolución de las bandas de las imágenes de Sentinel-2.

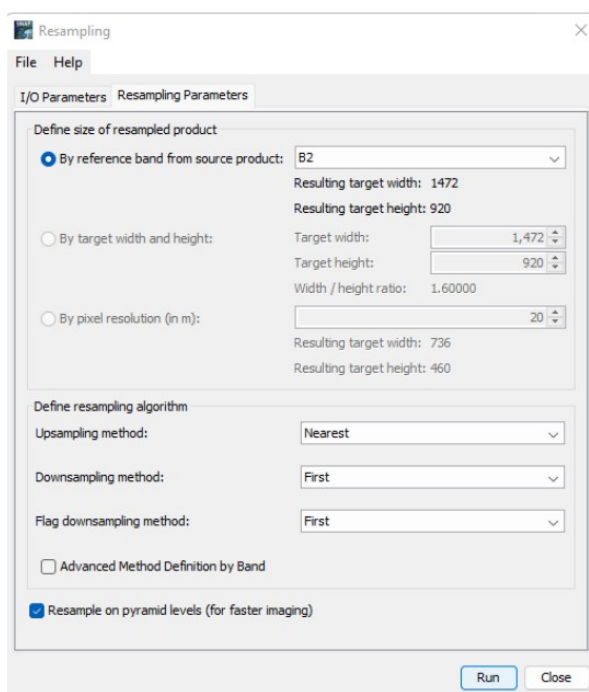


Figura 6: Ventana de Resampling.

En las figuras 7 y 8 se muestran las imágenes de la zona en 2022 y 2017, tras el proceso seguido de recorte y remuestreo.



Figura 7: Imagen de la zona de Orán-Es Senia en 2022



Figura 8: Imagen de la zona de Orán-Es Senia de 2015.

5.3.1.2 Pre-procesamiento de imágenes de Sentinel-1

La preparación de las imágenes de Sentinel-1, por su particularidad de ser capturadas con un instrumento radar (SAR), precisan de un procedimiento complejo con varios pasos que se describen en los apartados siguientes y se resumen en la figura 9. El objetivo es obtener una imagen de coherencia entre las dos imágenes de radar utilizadas para discriminar las áreas urbanas (con alta coherencia) de las áreas no urbanas, es decir, cultivos y vegetación natural (con baja coherencia) y en su caso elaborar una máscara de las zonas urbanas.

En este caso, por la complejidad de las operaciones realizadas, se ha recurrido al uso del constructor de gráficos (Grapho Builder). Esta herramienta permite al usuario ensamblar gráficos a partir de una lista de operadores disponibles y conectar nodos de operadores a sus fuentes.

A medida que se agregan operadores, sus interfaces de usuario correspondientes se crean y agregan como pestañas a una hoja de propiedades. Los operadores aceptan la entrada del usuario para la configuración del operador. En el caso del pre-procesamiento se ha seguido el siguiente diagrama de flujo (Figura 9):



Figura 9: Diagrama de flujo del pre-procesamiento de las imágenes de Sentinel-1

5.3.1.2.1 Tops-Split

Cada banda ancha interferométrica (IW) se compone de 3 sub-bandas y cada sub-banda tiene un máximo de 9 franjas. En SNAP, sólo se puede procesar una hilera a la vez hasta la etapa de Deburst. Nuestra zona de interés se encuentra en la franja IW2 y está suficientemente cubierta por el procesamiento de 2 ráfagas. Utilizaremos el operador TOPSAR-Split; así reduciremos el tiempo total de procesamiento (Figura 10).

Para añadir el operador TOPSAR-Split en Graph Builder se utiliza:

Add → Radar → Sentinel-1 TOPS → TOPSAR-Split.

En la pestaña TOPSAR-Split, ampliar el producto y seleccionar: Subswath: IW3, Polarisation: VV, Ráfagas: 4 y 5 (arrastra los dos deslizadores en consecuencia).

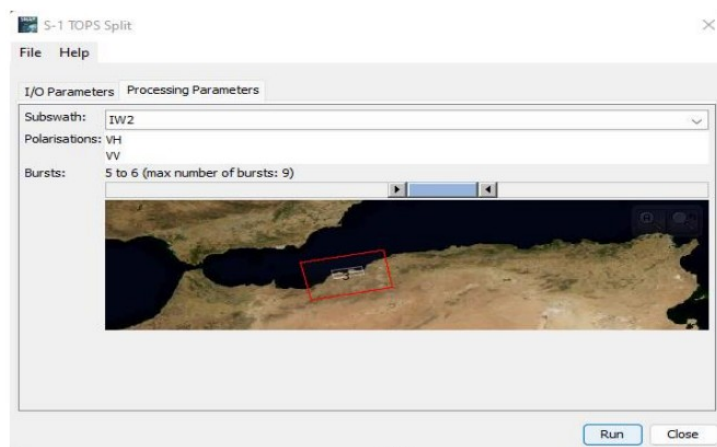


Figura 10. Ventana de TOPSAR Split



5.3.1.2.2 Apply-Orbit-File

A continuación, aplicaremos las órbitas actualizadas a los productos, añadiendo la herramienta:

Add → Radar → Aplicar archivo de órbita.

En este caso, se repite la misma operación para añadir el operador Apply-Orbit-File(2) y conectarlo al operador TOPSAR-Split(2).

Los vectores del estado de la órbita proporcionados en los metadatos de un producto SAR no suelen ser precisos y pueden refinarse utilizando los archivos de órbita precisos que están disponibles entre unos días y unas semanas después de la generación del producto. El archivo de órbita proporciona información precisa sobre la posición y la velocidad del satélite. A partir de esta información, se actualizan los vectores de estado de la órbita en los metadatos abstractos del producto.

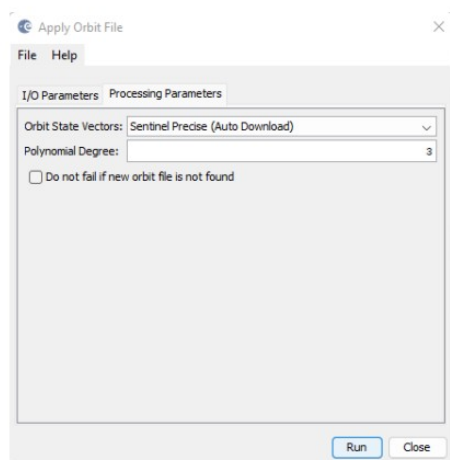


Figura 11: Ventana de Apply Orbit File

5.3.1.2.3 Create stack

Esta operación permite crear un apilamiento de imágenes. Para ello se añade la herramienta:

Add → Radar → Coregistration → Stack tools → Create stack.

Sólo tenemos que mantener los dos últimos. Seleccione todos los productos anteriores, para la creación de un stack, la información relacionada aparecerá en los otros campos como se muestra a continuación.

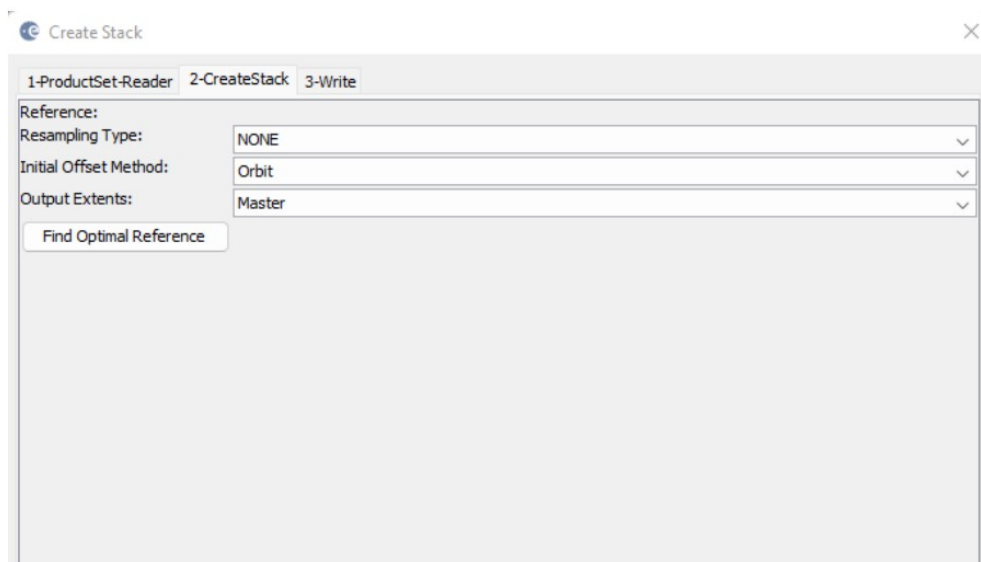


Figura 12: Ventana de Create Stack

5.3.1.2.4 Back geocoding

Ahora vamos a co-registrar los dos productos de S-1 SLC de la misma sub-banda obtenidos mediante el paso "aplicar archivo de órbita", utilizando las órbitas de ambos productos y un modelo digital de elevación (DEM), y para ello se utiliza la herramienta :

Add -> Radar -> Corregistro -> Corregistro S1 TOPS -> Back-Geocoding. Defina los dos operadores Aplicar archivo de órbita como entrada.

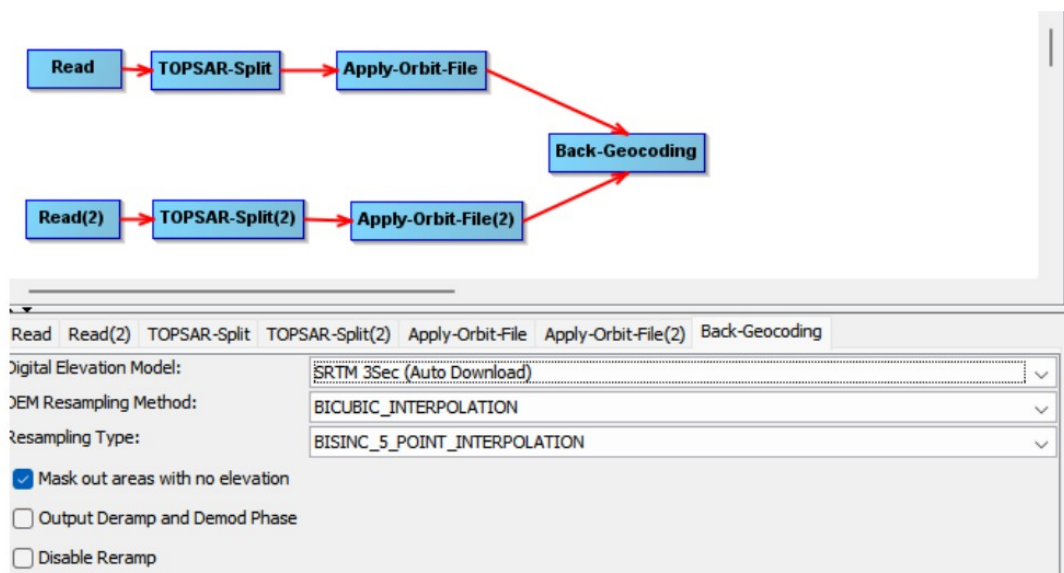


Figura 13: Ventana de Back Geocoding



5.3.1.2.5 Enhanced Spectral Diversity

Este operador sigue el operador de geocodificación inversa, estimando primero un desplazamiento de alcance constante para cada ráfaga utilizando un pequeño bloque de datos en el centro de la ráfaga, y luego estimando un desplazamiento de acimut constante. Por último, se promedian las estimaciones de todas las ráfagas para obtener el alcance y el desplazamiento acimutal constantes finales de toda la imagen. Para ello se añade la herramienta:

Add → Radar → Coregistration → S1 TOPS Coregistration → S1 TOPS Coregistration.

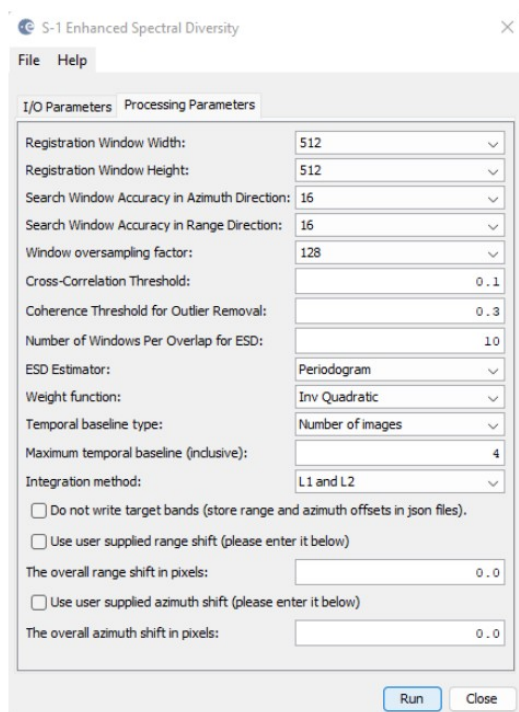


Figura 14: Ventana de Enhanced Spectral Diversity

5.3.1.2.6 Estimación de la coherencia

Una vez tenemos en el producto corregido las bandas de intensidad de ambas imágenes (2019 y 2022), utilizaremos el operador de Coherencia, para estimar lo bien que están "correlacionados" los píxeles entre las imágenes maestra y secundaria utilizadas.

Add → Radar → Interférométrie → Produits → Cohérence

En la pestaña Coherencia, el tamaño de la ventana de rango de coherencia se establece en 20 y el tamaño de la ventana de azimuth de coherencia se establece automáticamente en 5.

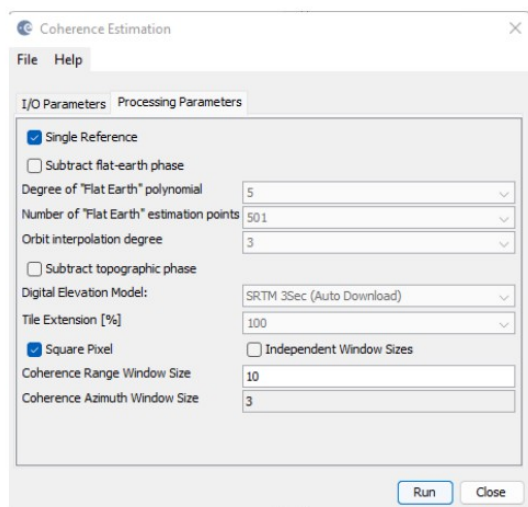


Figura 15: Ventana de Coherence Estimation

5.3.1.2.7 TOPS Deburst

Esta operación se realiza para eliminar el "espacio negro" entre las dos ráfagas. Se aplica con:

Add → Radar → Sentinel-1 TOPS → TOPSAR-Deburst

En la pestaña TOPSAR-Deburst, se mantiene la configuración por defecto (Polarisations: VV).

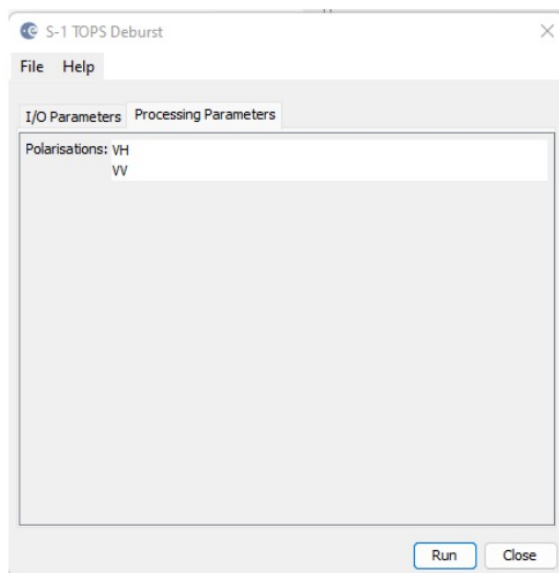


Figura 16: Ventana de TOPS Deburst

5.3.1.2.8 Multilooking

Aplicando este operador, reduciremos el ruido de moteado inherente que aparece inicialmente en las imágenes SAR y obtendremos píxeles cuadrados. Se aplica con la opción:

Add → Radar → Utilities SAR → Multilooking



En la pestaña Multilook, se mantiene seleccionada la opción "Píxel cuadrado GR" y se ajusta el número de ojos de alcance a 8. El número de observaciones de acimut se aumentará a 2 y el píxel cuadrado GR medio a 27,348389.

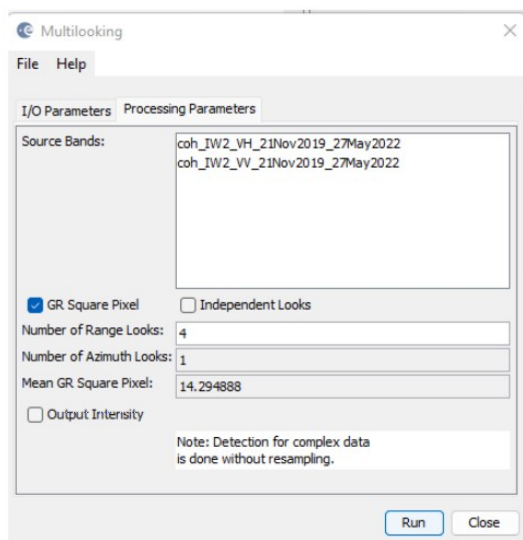


Figura 17: Ventana de Multilooking

5.3.1.2.9 Corrección del terreno

Se aplica el operador de corrección del terreno, para convertir las coordenadas del RADAR en coordenadas geográficas, lo que se hace con la herramienta:

Radar → Geometric → Terrain correction → Rangle Doppler Terrain-Correction



Identificación de barrios de infraviviendas (*slums*) en Argelia mediante teledetección

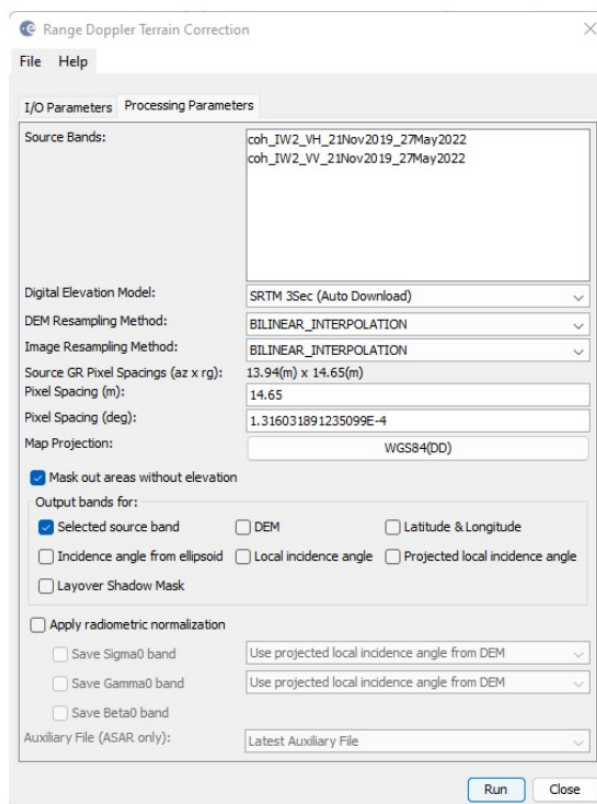


Figura 18: Ventana de Terrain Correction

La imagen de coherencia filtrada y corregida geoméricamente se muestra en la figura 18. Se aprecia la diferencia entre las zonas de alta coherencia (urbanas, más claras) y las zonas de baja coherencia (no urbanas, más oscuras).

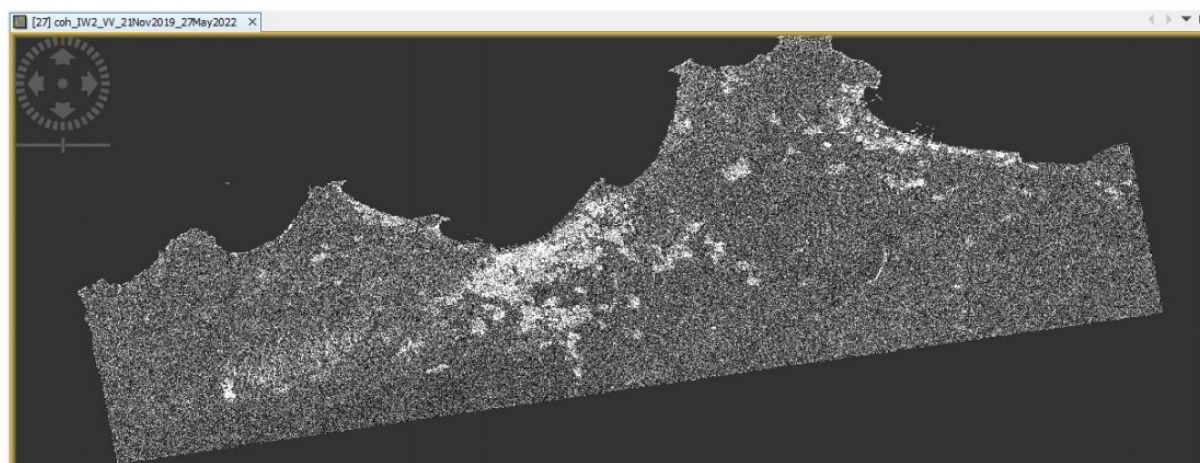


Figura 19. Imagen de coherencia entre las dos imágenes de radar de Sentinel-1



5.3.2 Clasificación

La etapa de clasificación implica un proceso para asignar finalmente cada píxel u objeto de la imagen a una clase de carácter cualitativo, como la cobertura del suelo. Se trata, pues, de un proceso de generalización que permite pasar de un conjunto de valores cuantitativos, los niveles digitales de cada banda de la imagen, a un valor cualitativo, la clase a la que se asigna cada píxel. El proceso consiste en clasificar los píxeles de una imagen en diferentes clases en función de sus características

Los algoritmos de clasificación son muy diversos, pero pueden dividirse en dos categorías principales: la clasificación supervisada, que utiliza un conjunto de datos entrenamiento para calibrar el algoritmo de antemano; y la clasificación no supervisada, que no realiza un entrenamiento a priori, sino que realiza la clasificación partiendo únicamente de los datos de la imagen y estableciendo agrupaciones o cluster a partir de los algoritmos.

En la clasificación supervisada se han desarrollado métodos basados en la estadística clásica multivariante como los métodos de mínima distancia, máxima probabilidad, análisis discriminante, etc.; más recientemente se han desarrollado y aplicado los métodos de aprendizaje automático, como los árboles decisión (*classification and regression trees*, CART), los bosques aleatorios (*random forest*), *k* vecino más próximos (*k-nearest neighbours*, *k*-NN), máquinas de soporte vectorial (*support vector machines*, SVM), entre otros; y dentro de ellos el grupo de las redes neuronales desde las más sencillas de tipo perceptrón y perceptrón multicapa (*multilayer perceptron*, MLP) hasta las más complejas como las redes neuronales convolucionales (*convolutional neural networks*, CNN), de gran aplicación en el análisis de imágenes. Por su parte, dentro de los métodos no supervisados hay también varios como el conocido algoritmo *k*-medias (*k-means*), esperanza-maximización (*expectation-maximization*, EM) o ISODATA. Algunos de ellos están implementados en SNAP.

Para este trabajo se va emplear el método supervisado de los bosques aleatorios, más conocido por el término inglés de *random forest* (Breiman, 2000). Se trata de un método basado en los árboles de decisión, pero utilizando un número elevado de árboles, lo que compensa los errores en la decisión que introduce el método convencional de árboles. Está dentro de las denominadas técnicas de ensamblaje de modelos, que consisten en agregar un gran número de modelos de prueba para mejorar la precisión de la clasificación y la regresión. Es un método ampliamente utilizado que en general da un buen rendimiento.

5.3.2.1 Clasificación de imágenes de Sentinel-2

El flujo de operaciones realizados en Sentinel-2 para la clasificación y la obtención de la máscara urbana se ha mostrado en la Figura 4, junto al pre-procesamiento tratado en títulos anteriores (*preprocesamiento de imágenes de Sentinel-2*).

La clasificación supervisada se realiza en dos etapas principales. La primera etapa se denomina etapa de entrenamiento y la segunda etapa se denomina etapa de clasificación. Por lo tanto, previamente a la ejecución de la clasificación supervisada hay que realizar el entrenamiento, que en SNAP se hace a través de la digitalización o importación de los vectores (polígonos) de entrenamiento (training vectors) de cada una de las clases que van a ser diferenciadas. Así, el usuario elige un número de vectores (polígonos, rectángulos, elipses), denominado muestra de entrenamiento, cada uno de los cuales representa una clase distinta, como agua, urbano, bosque matorral, cultivos, etc. a través del cual se realiza la clasificación supervisada. Las regiones pueden crearse utilizando la herramienta "Nuevo Vector Data Container" y otras herramientas de dibujo como la "Herramienta de dibujo de rectángulos", como se muestra en la Figura 20.

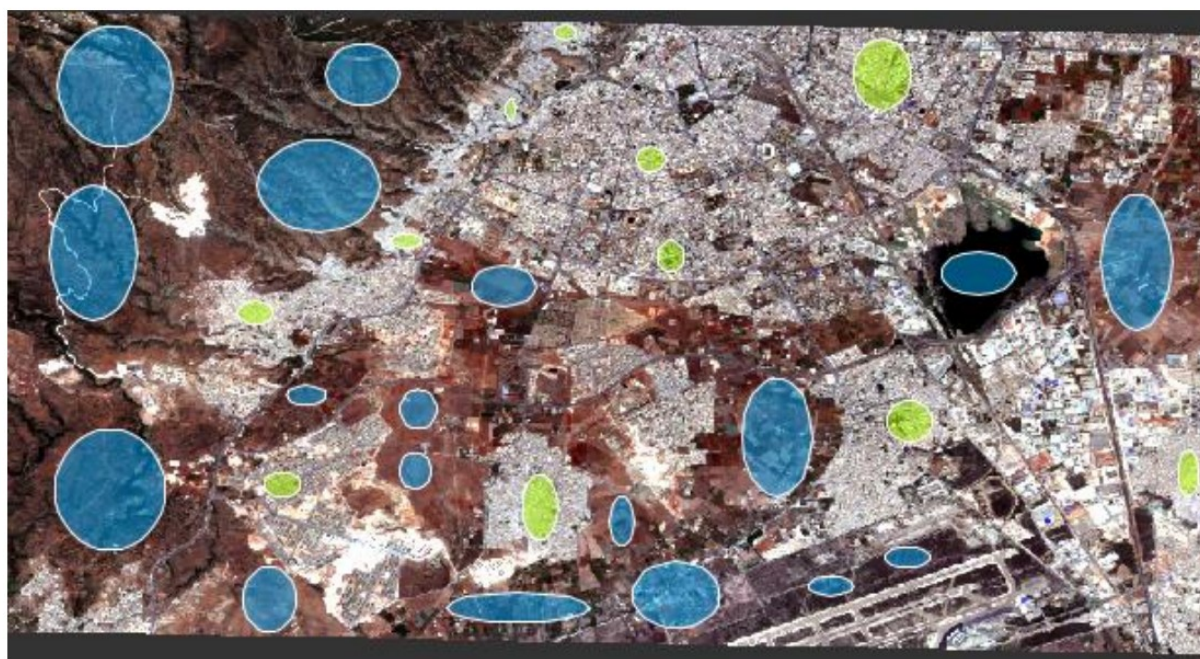


Figura 20: Definición de los polígonos o vectores de entrenamiento.

Una vez realizado el entrenamiento, se pasa a la segunda fase que es la clasificación propiamente dicha, mediante el algoritmo de random forest. Para ello se introduce la imagen a clasificar (Figura 21) y luego se definen los parámetros que se emplean en la clasificación (Figura 22), número de



muestras de entrenamiento (5000 por defecto), número de arboles (10 por lo menos), las muestras de entrenamiento (tantas como clases) y las bandas utilizadas (en principio todas las de Sentinel-2)

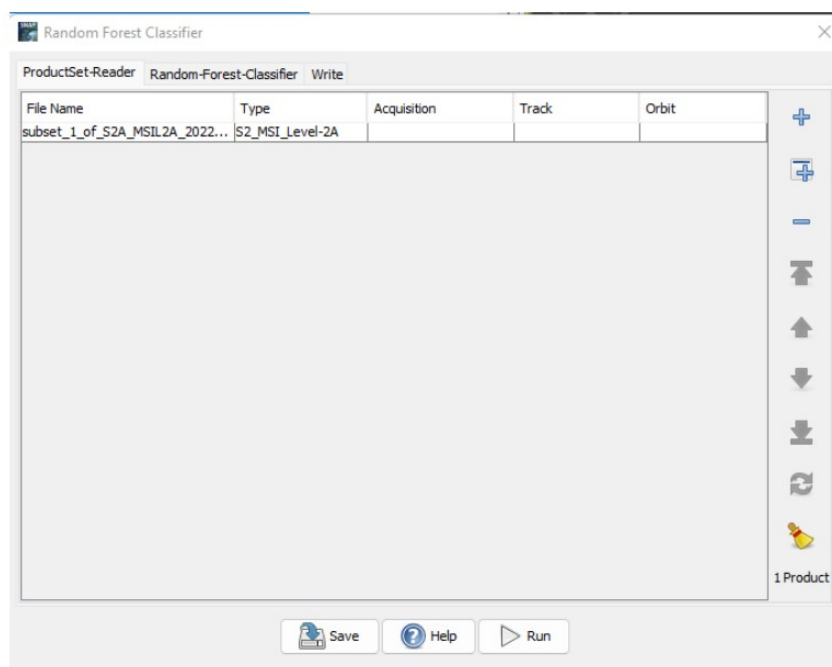


Figura 21: Ventana de clasificación random forest (entrada de datos)

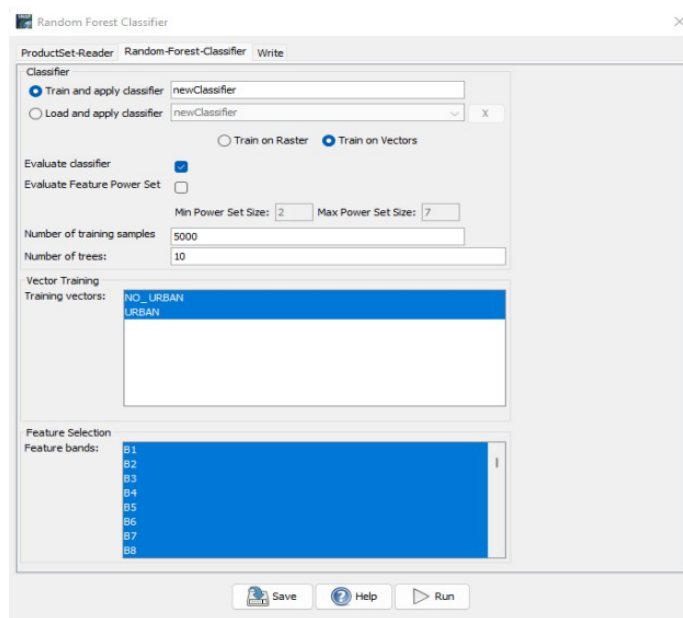


Figura 22: Ventana de clasificación random forest (parámetros)

Se ha tratado de hacer una primera clasificación con 8 clases en la zona de Es Senia: urbano respetable, nuevas edificaciones, slums, rutas, suelos, cultivos, prados y bosques (Figura 23). La clasificación distingue bien entre los usos de suelo urbanos y el resto, pero no discrimina bien los slums de otros

tipos urbanos como la vivienda respetable y las nuevas edificaciones, por lo que adopta una nueva estrategia. No obstante, puede ser utilizada como una clasificación de cobertura del suelo.

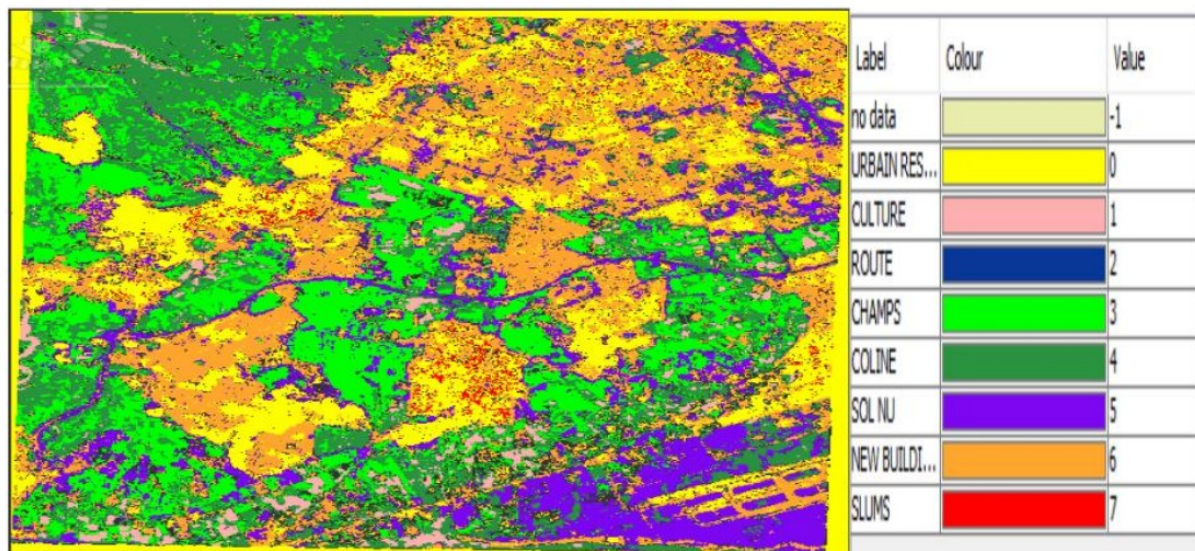


Figura 23: Clasificación general de las coberturas del suelo mediante random forest en la zona de Es Senia.

La estrategia diseñada es emplear el análisis textural (Wurm et al., 2017) ya que la diferencia entre los slums y otros tipos urbanos es más bien textural que espectral, lo que explica que la clasificación inicial no funcione correctamente. Sin embargo, la aplicación directa del análisis textural sobre toda la imagen puede enmascarar o dificultar el objetivo que es distinguir entre distintas estructuras urbanas. Por ello se ha aplicado una máscara de zonas urbanas sobre una banda de la imagen completa (la banda 8), para lo que es necesario obtener la máscara.

La máscara se obtiene a partir de una clasificación en la que se distinguen las coberturas anteriores (suelos, cultivos, prados y bosques) más un solo tipo urbano. A partir de ahí con la herramienta de Band Maths (Figura 24), se aplica una reclasificación que asigna 0 a todas las clases excepto a la clase urbana a la que se asigna 1. La expresión utilizada (habiendo asignado originalmente 0 a la clase urbana) es:

```
'if$4.LabeledClasses == 0 then1 else0'
```

Que discrimina entre la zona urbana (1) y la zona no urbana (0), como se muestra en la Figura 25.

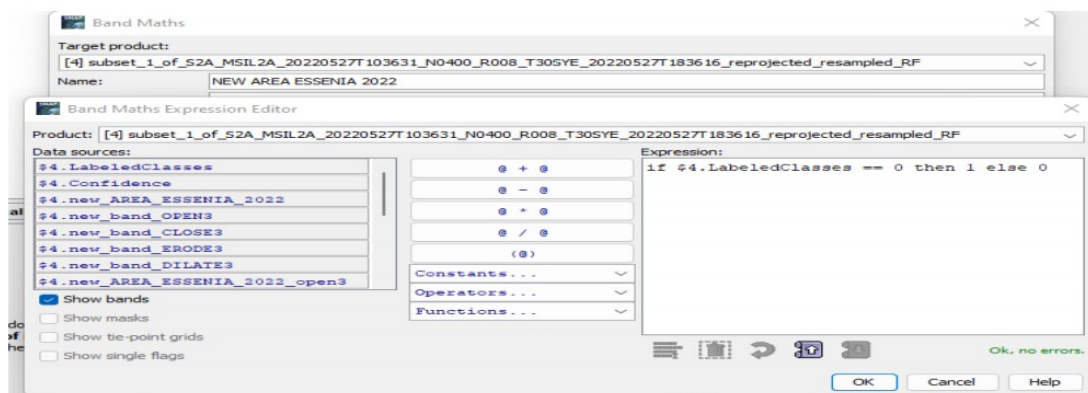


Figura 24: Ventana de Band maths, donde se aplica la reclasificación

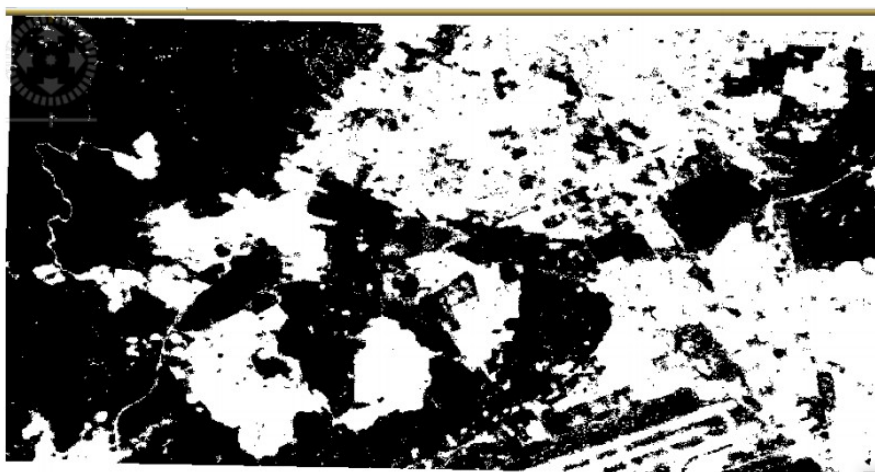


Figura 25: Máscara de la zona urbana según Sentinel-2 (1: Urbano; 0: No urbano)

La máscara obtenida se aplica sobre la imagen, en concreto sobre la banda 8 (infrarrojo próximo) que suele ser una banda utilizada en este tipo de estudios, al proporcionar una buena información para diferenciar distintos tipos de cubiertas. Hay que tener en cuenta que los análisis texturales se aplican habitualmente sobre una sola banda, generalmente la que proporcione una mayor información. En este caso como se ha indicado se ha utilizado la banda 8, aunque se podrían utilizar otras bandas (las del visible que presentan una mayor resolución original en las imágenes de Sentinel-2 o incluso alguna componente principal, si se aplicara este análisis). La aplicación de la máscara sobre la banda 8 original se hace nuevamente mediante Band Maths, utilizando una multiplicación de ambas (Figura 26):

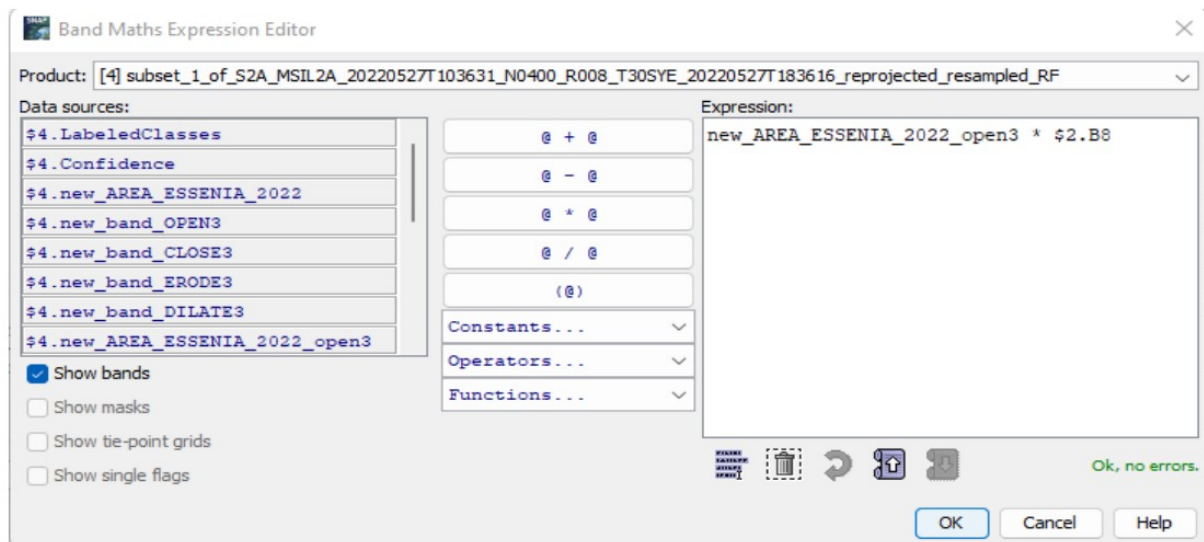


Figura 26: Ventana de Band Maths, donde se aplica la máscara con una operación de multiplicación

El resultado se muestra en la Figura 27 , donde se aprecia la banda 8 restringida a las zonas urbanas en 2022 .

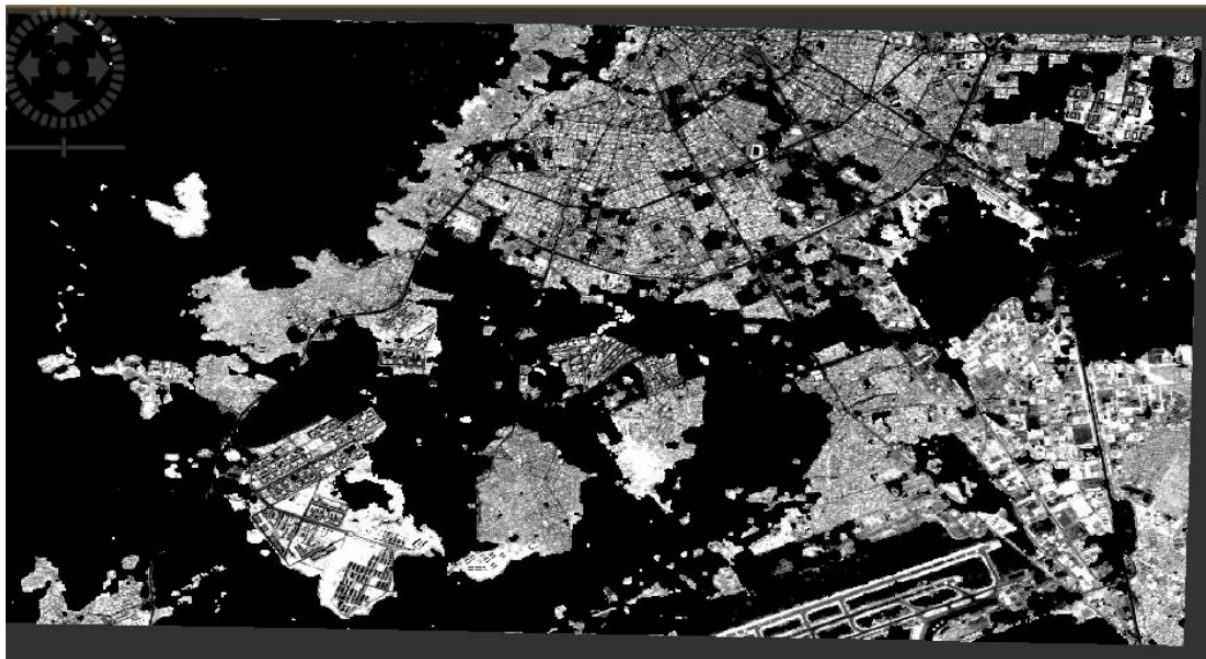


Figura 27: Banda 8 de la imagen Sentinel-2 enmascarada en las zonas urbanas en 2022.

5.3.2.2 Clasificación de imágenes de Sentinel-1

La clasificación de la imagen de Sentinel 1 se lleva a cabo mediante random forest sobre la imagen.

Para este paso, utilizaremos de nuevo el algoritmo de bosque aleatorio. Para ello, introducimos la imagen a clasificar (Figura 19), la de la coherencia, y a continuación definimos los parámetros de la misma forma que la utilizada en el paso anterior para la clasificación del centinela 2, el número de

muestras de entrenamiento (5000 por defecto), el número de árboles (al menos 10), las muestras de entrenamiento (tantas como el número de clases) y las bandas utilizadas.

El resultado de la clasificación se muestra en la Figura 28.

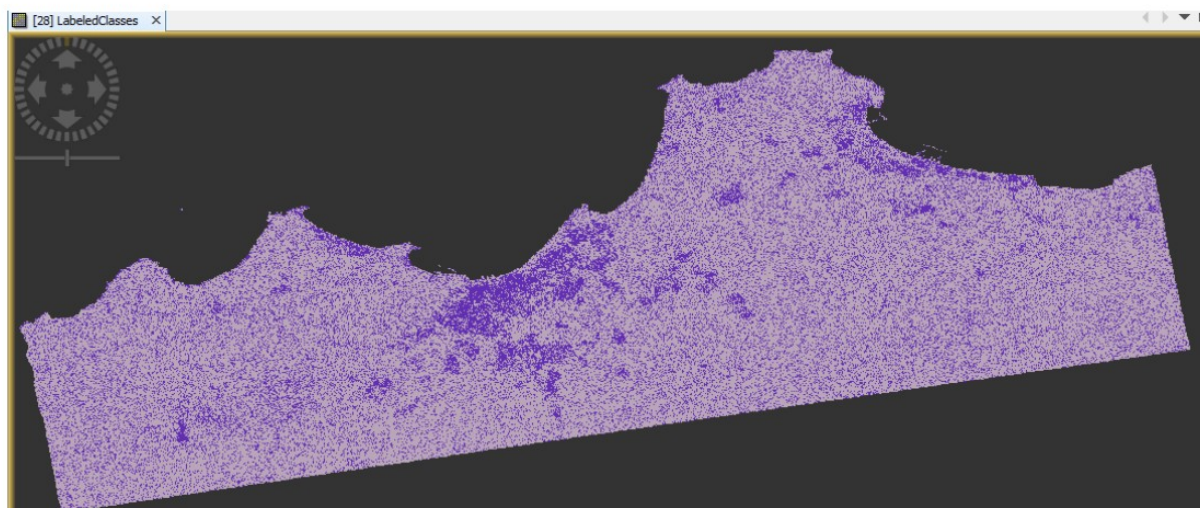


Figura 28: Imagen de Sentinel-1 clasificada con random forest

Una obtenida la clasificación se obtiene la máscara con Band Maths, como anteriormente, mediante una expresión condicional que permite reclasificar la imagen clasificada:

```
'if$4.LabeledClasses == 0 then1 else 0
```

Que nos permite determinar entre la zona urbana (=1) y la zona no urbana (=0), como se muestra en la Figura 29.

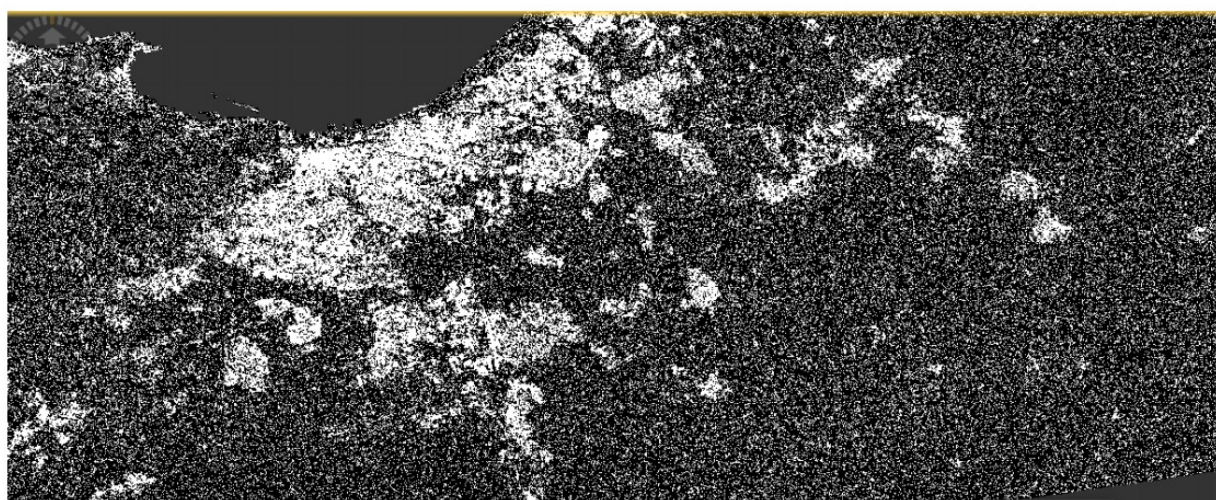


Figura 29: Máscara de la zona urbana según Sentinel-1 (1: Urbano; 0; No urbano)

Sin embargo, la máscara obtenida mediante imágenes Sentinel-1 muestra mucho más ruido que la obtenida con Sentinel-2, y por lo tanto su eficacia como máscara es mucho menor, sobre todo si pretendemos aplicarla sobre una imagen original con el objetivo de detectar texturas para identificar los slums. Por ello, a partir de esta fase, continuaremos nuestro trabajo con las imágenes Sentinel-2.

5.3.3 Aplicación de filtros texturales

La detección de slum, discriminando de otros tipos de zonas urbanas en la imagen recortada por la máscara, se realiza mediante la aplicación de filtros texturales (análisis textural), concretamente filtros de co-ocurrencia de nivel de gris (*Grey Level Co-occurrence Matrix*, GLCM), aunque estará apoyada con operaciones de morfología matemática (previos y posteriores al análisis textural).

El diagrama de flujo de esta fase se muestra en la Figura 30.

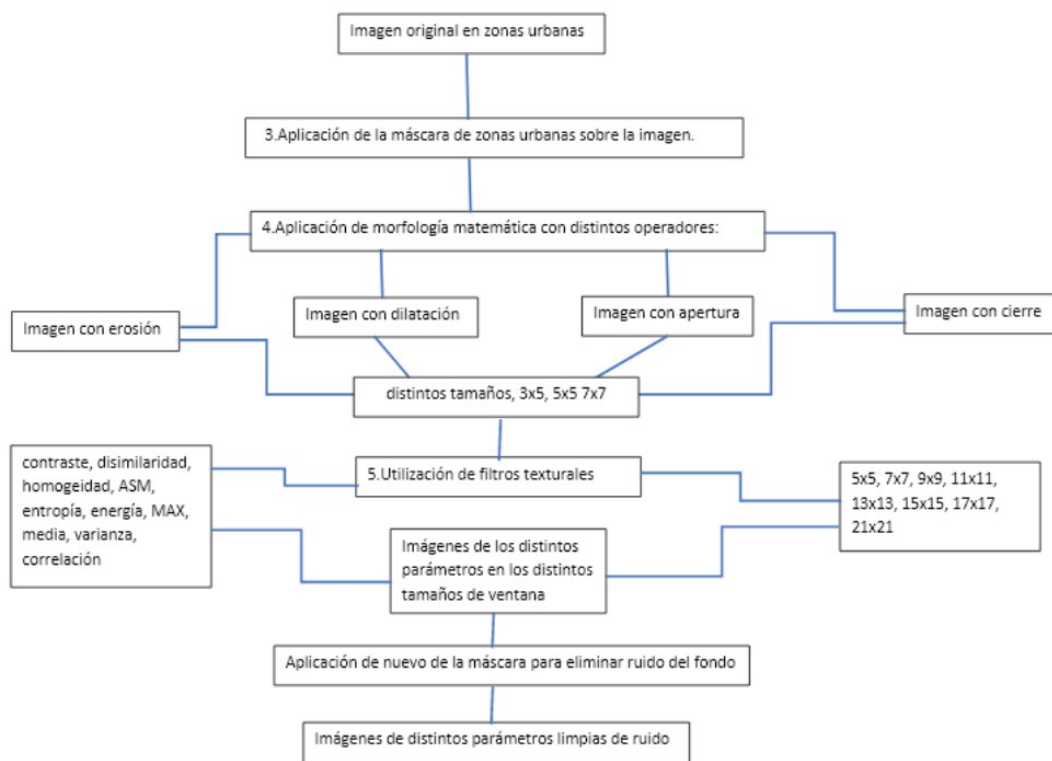


Figura 30: Diagrama de flujo del análisis textural y morfología matemática

5.3.3.1 Operadores de morfología matemática (pre-proceso)

Las operaciones de morfología matemática se aplican en SNAP sobre una banda determinada (botón derecho → Filtered Band) mediante matrices de filtrado en sentido estricto en las que el valor de un píxel objetivo se calcula a partir de la vecindad de $N \times M$ píxeles de un píxel fuente correspondiente.

Los operadores de morfología matemática son obviamente funciones no lineales en los que se utiliza un núcleo booleano $N \times M$, que actúa como elemento estructurador del filtro.

Una vez abierta la venta de filtros, se puede elegir dentro de los filtros morfológico qué filtro queremos utilizar y con qué tamaño de ventana, para tener obtener mejores resultados; sin embargo, en nuestro caso utilizaremos los diferentes filtros y así podremos elegir el mejor de ellos. Como operadores se dispone de la erosión, dilatación, apertura y cierre, para ventanas de 3x3, 5x5 y 7x7.

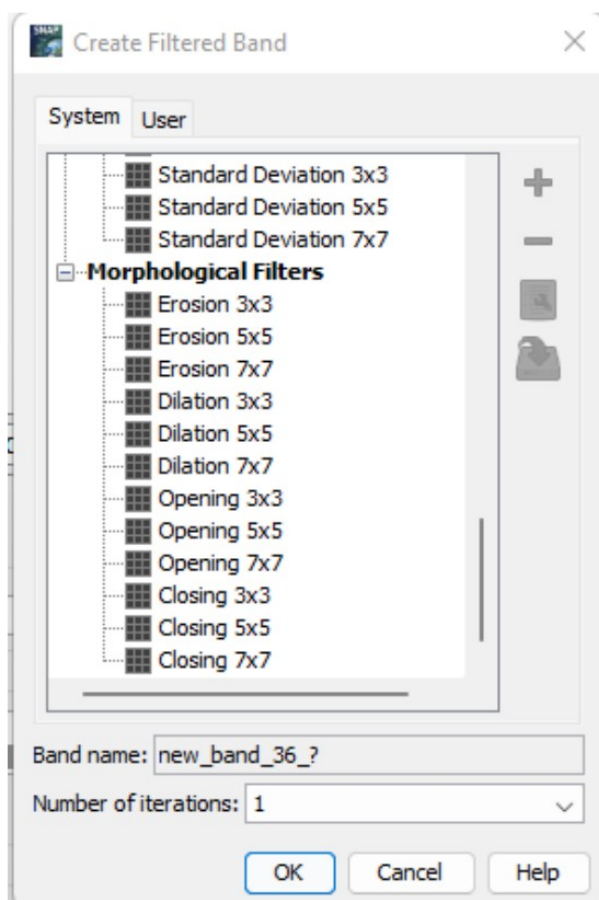


Figura 31: Ventana de los Filtros morfológicos

5.3.3.1.1 Dilatación

Una dilatación morfológica consiste en desplazar el elemento estructurante sobre cada píxel de la imagen, y observar si el elemento estructurante "toca" (o más formalmente interseca) la estructura de interés. El resultado es una estructura más grande que la original. Dependiendo del tamaño del elemento estructurante, algunas partículas pueden conectarse y algunos agujeros pueden desaparecer. En la figura 32 se muestran los resultados de la aplicación de ese filtro.

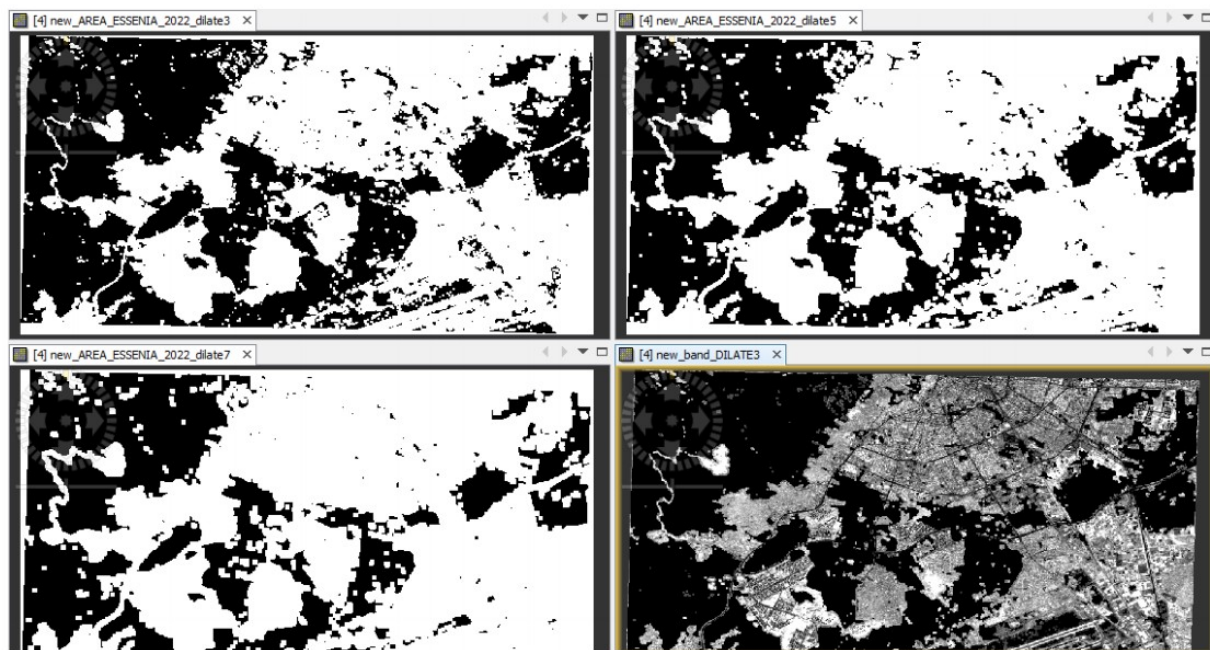


Figura 32: Resultado de aplicar el filtro de dilatación sobre la máscara y la imagen en la zona de Es Senia.

5.3.3.1.2 Erosión

La erosión es la operación opuesta, y se define como una dilatación del complemento de la estructura. Consiste en buscar todos los píxeles para los que el elemento estructurante centrado en ese píxel toca el exterior de la estructura. El resultado es una estructura erosionada o recortada, como se muestra en la Figura 33. Se observa la desaparición de las partículas más pequeñas que el elemento estructurante utilizado y la posible separación de las partículas grandes.

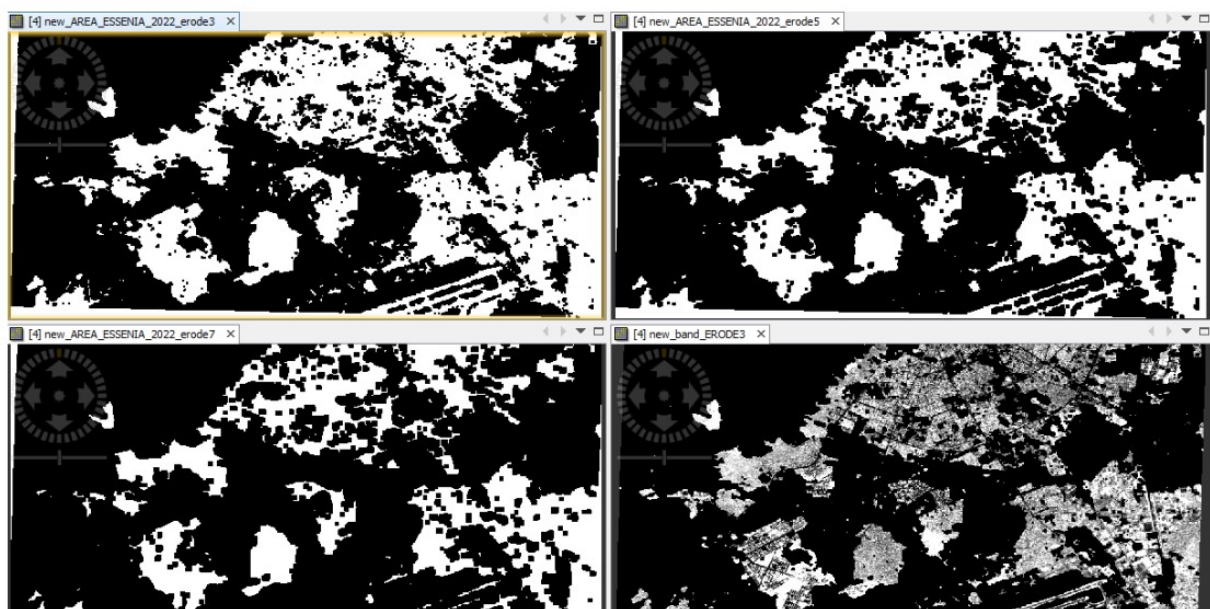


Figura 33: Resultado de aplicar la erosión en la imagen y la máscara en la zona de Es Senia.

La erosión y la dilatación tienen la desventaja de cambiar fuertemente el tamaño de las estructuras en la imagen. Para reducir este efecto, a menudo se utilizan en combinación. Así, la apertura y el cierre morfológicos tienen una propiedad de idempotencia: el resultado no cambia si se aplica el operador varias veces, basta con aplicarlo una sola vez.

La apertura y el cierre morfológico cambian relativamente poco la forma de las partículas grandes. Por otro lado, permite eliminar fácilmente pequeñas partículas aisladas o pequeños agujeros dentro de las estructuras. Por lo tanto, suelen utilizarse para limpiar el resultado de una binarización.

5.3.3.1.3 Apertura

La apertura morfológica (Figura 34) se define como una erosión seguida de una dilatación. Los efectos de la apertura son:

1. Eliminar las partículas pequeñas (más pequeñas que el tamaño del elemento estructurante)
2. Separar las partículas grandes en los puntos donde son más finas.

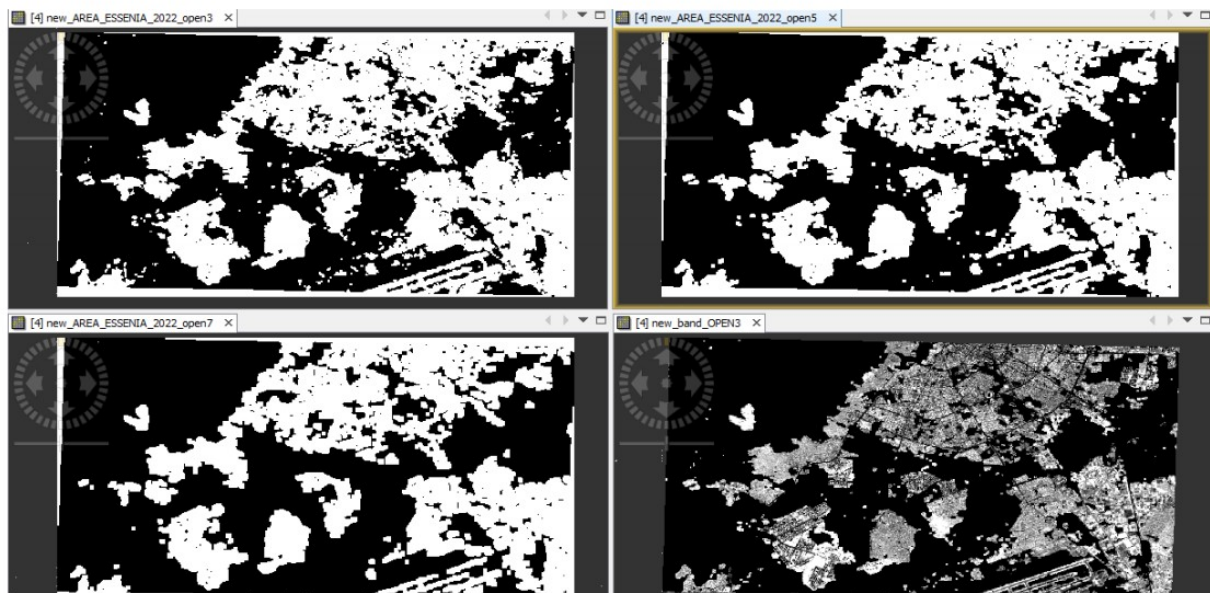


Figura 34: Resultado de aplicar la apertura en la imagen y la máscara en la zona de Es Senia

5.3.3.1.4 Cierre

Por su parte, el cierre morfológico (Figura 35) se define como una dilatación seguida de una erosión. El cierre tiene el efecto de eliminar los pequeños agujeros de las estructuras conectar las estructuras cercanas

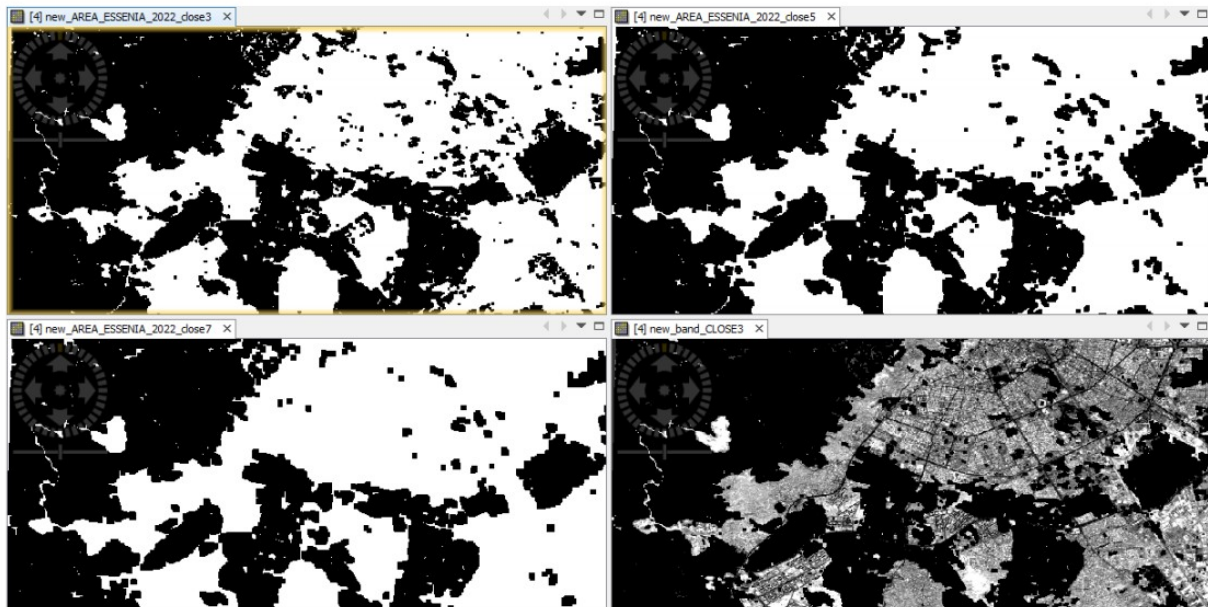


Figura 35: Resultado de aplicar el cierre en la imagen y la máscara en la zona de Es Senia.

Como resultado de la aplicación de los filtros morfológicos con los diferentes operadores así como los diferentes tamaños, podemos observar que se pueden eliminar muchos residuos de la imagen para tener una imagen clara y nítida en varios aspectos, por ejemplo, con el operador de apertura con el que vamos a continuar nuestro trabajo. También podemos tener buenos resultados con el operador de cierre; no así con los otros operadores, como la erosión y la dilatación que, no obstante, se pueden utilizar para otros fines.

5.3.3.2 Filtros texturales (Grey Level Co-occurrence Matrix): tipos y tamaño de ventana

La información espacial en forma de características de textura puede ser útil para la clasificación de imágenes. Las mediciones de textura pueden producir nuevas imágenes utilizando la información espacial inherente a la imagen. La textura es el patrón de variaciones de intensidad en una imagen y puede ser una herramienta valiosa para mejorar la precisión de la clasificación de la cobertura del suelo. La matriz de co-ocurrencia de niveles de gris (GLCM) es uno de los métodos más utilizados para calcular medidas de textura de segundo orden. Fueron definidos por primera vez por Haralick et al. (1973) y son muy utilizados en el análisis textural a partir de entonces.

A partir de la matriz GLCM pueden calcularse varias características de textura, por ejemplo, el segundo momento angular, el contraste, la correlación, la entropía, la varianza, el momento diferencial inverso, la media de las diferencias, la varianza de las diferencias, la entropía de las diferencias, la media de las sumas, la varianza de las sumas y la entropía de las sumas.



Cada característica modela diferentes propiedades de la relación estadística de la co-ocurrencia de píxeles estimada dentro de una ventana móvil dada y a lo largo de direcciones predefinidas y distancias entre píxeles. GLCM es una medida de la probabilidad de ocurrencia de dos niveles de gris separados por una distancia dada en una dirección determinada. Los disponibles en SNAP son los siguientes (Figura 36):

- Contraste (Contrast)
- Disimilitud (Dissimilarity, DIS)
- Homogeneidad (Homogeneity, HOM)
- Segundo Momento Angular (Angular Second Moment, ASM)
- Energía (Energy)
- Máxima Probabilidad (Maximum Probability, MAX)
- Entropía (Entropy, ENT)
- Media (GLCM Mean)
- Varianza (GLCM Variance)
- Correlación (GLCM Correlation)

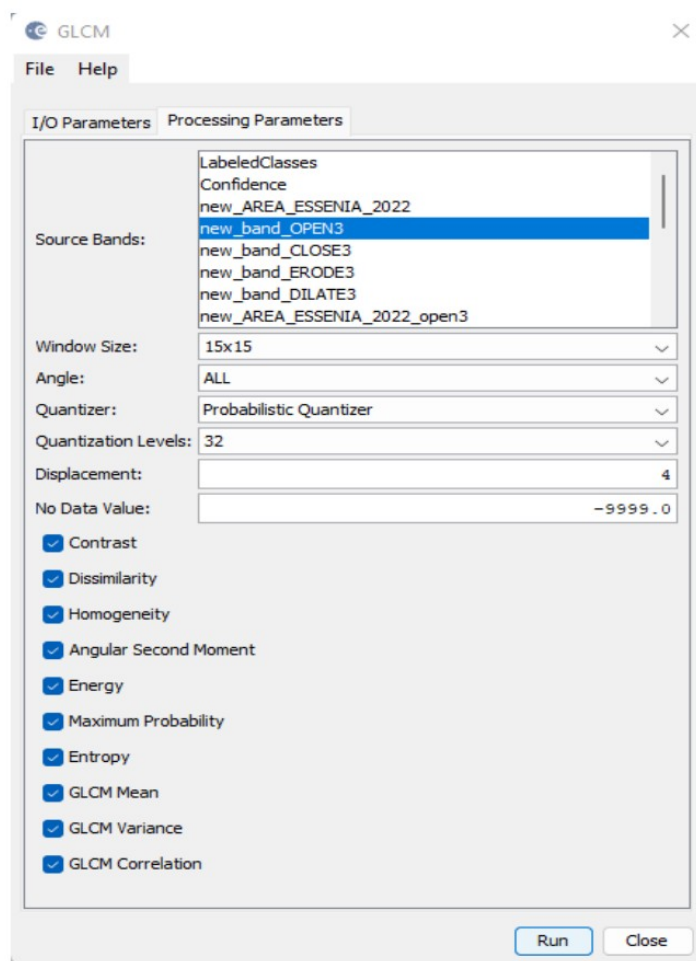


Figura 36: Ventana de filtros texturales GLCM

En este trabajo se han calculado todos los filtros texturales GLCM (contraste, disimilitud, homogeneidad, ASM, entropía, energía, MAX, media, varianza, correlación), con diferentes tamaños de ventana (5x5, 7x7, 9x9, 11x11, 13x13, 15x15, 17x17, 21x21); para, mediante análisis visual, seleccionar aquellos que resultan más eficaces para el objetivo propuesto de identificación de barrios de infravivienda o *slums*. Los filtros se aplicaron sobre la imagen de la banda 8 de Sentinel-2, recortada por la máscara de las zonas urbanas y tras la aplicación de los operadores de apertura y cierre, antes comentados.

En las figuras 37 a 39 se muestra el resultado de las imágenes resultantes después de aplicar distintos filtros texturales (GLCM) con diferentes tamaños de ventana. En concreto, en las imágenes se muestran los filtros de contraste, disimilitud, entropía (en escala de color) y varianza.

En cuanto al mejor resultado obtenido de los filtros, podemos deducir que la disimilitud nos da resultados más convincentes aunque también la entropía. Para los siguientes pasos se va a utilizar la disimilitud con una ventana de 15*15.

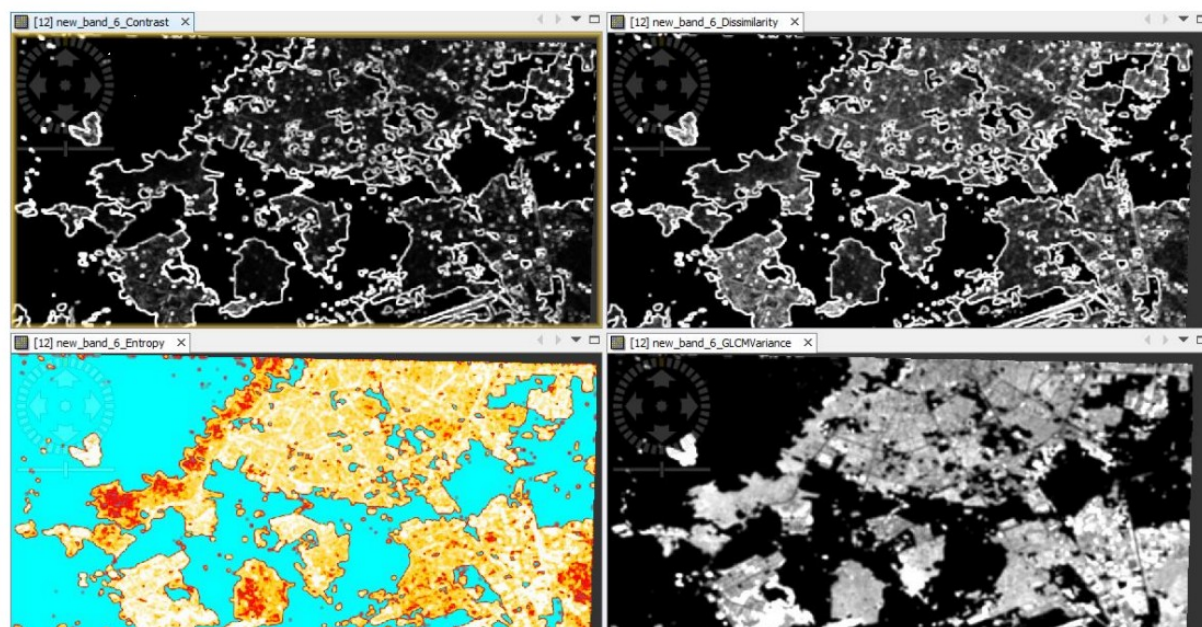


Figura 37: Resultados de la aplicación de filtros GLCM con matriz de 11x11 sobre la banda 8 en la zona piloto de Es Senia.

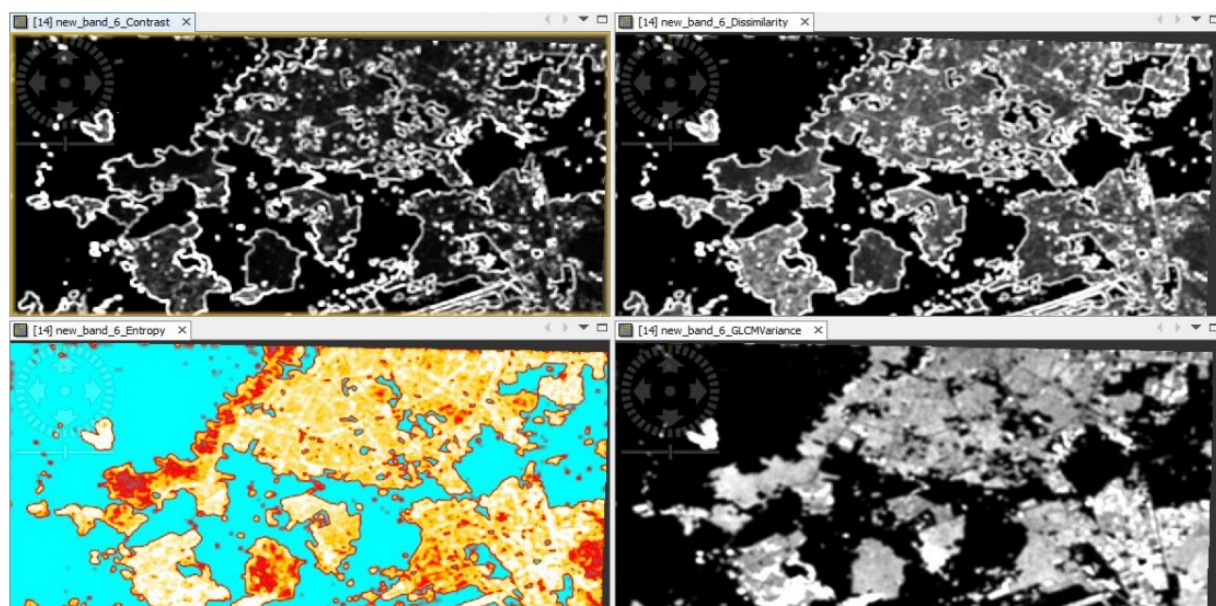


Figura 38: Resultados de la aplicación de filtros GLCM con matriz de 15x15 sobre la banda 8 en la zona piloto de Es Senia.

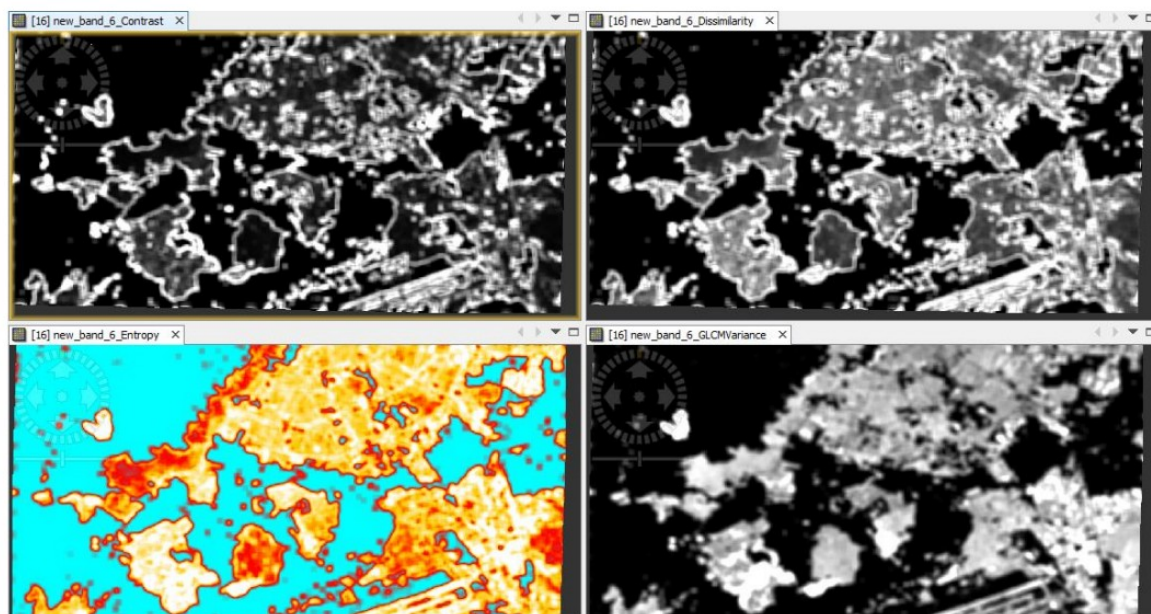


Figura 39: Resultados de la aplicación de filtros GLCM con matriz de 21x21 sobre la banda 8 en la zona piloto de Es Senia.

La representación de mapas de ciudades bajo diferentes filtros morfológicos.

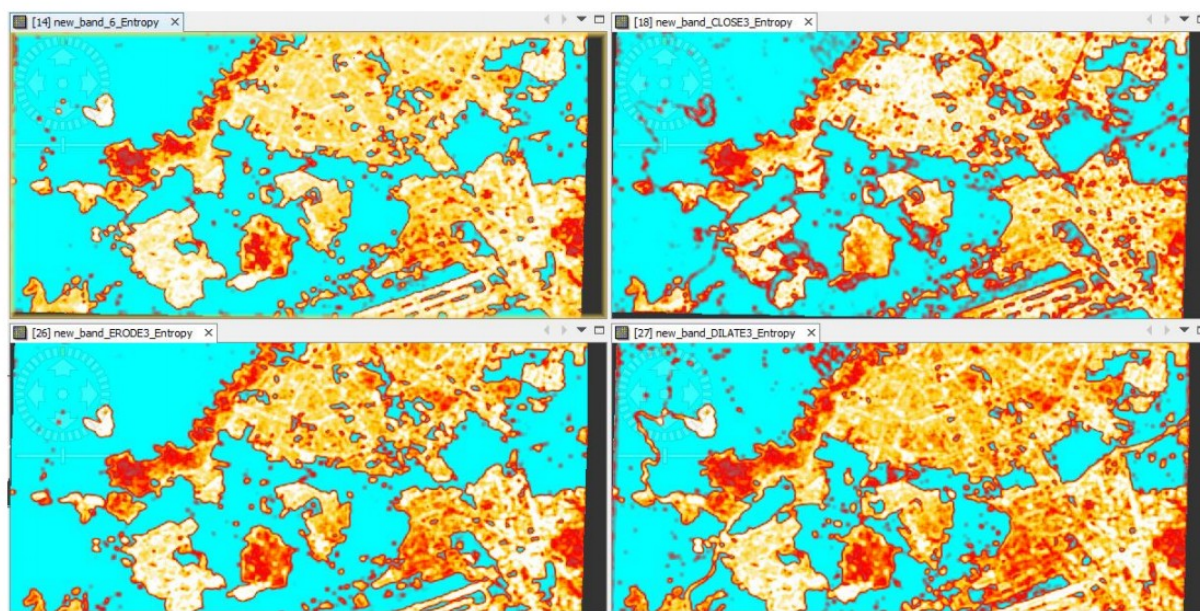


Figura 40: Resultados de la aplicación de filtros GLCM sobre imágenes con distinto pre-procesado en la zona piloto.

5.3.3.3 Operadores de morfología matemática (post-proceso)

Tras la aplicación de los distintos filtros texturales, con distintas ventanas y sobre imágenes sometidas a distintos operadores morfológicos (pre-procesamiento), se vuelven a aplicar operadores de morfología matemática (apertura) para mejorar los resultados, eliminando áreas residuales o reduciendo el grosor de bordes en las zonas urbanas. En las figuras 41 a 43 se muestran estos resultados.



Identificación de barrios de infraviviendas (*slums*) en Argelia mediante teledetección

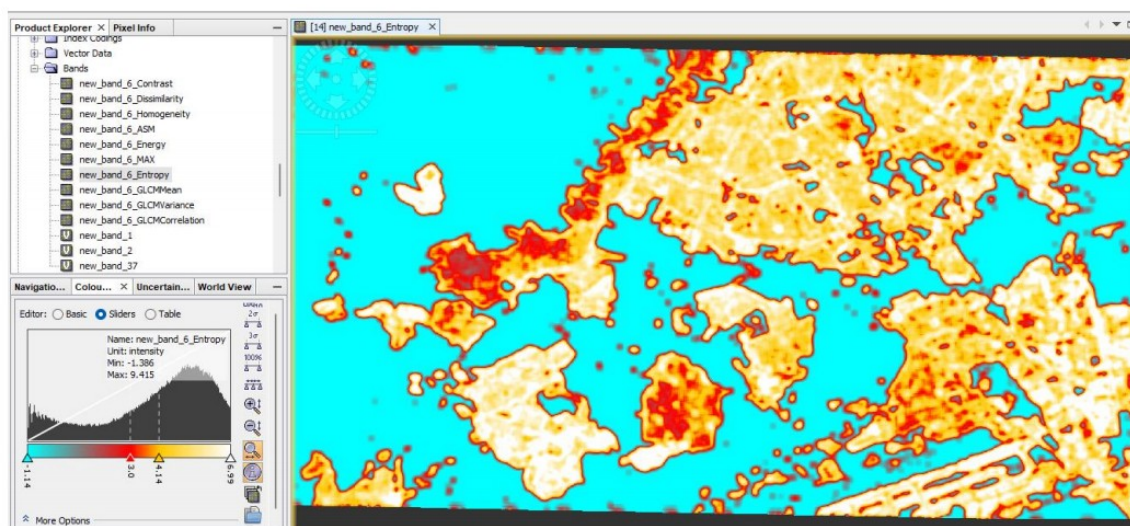


Figura 41: Resultados de la aplicacion de filtros GLCM sobre imagen de la entropia en la apertura

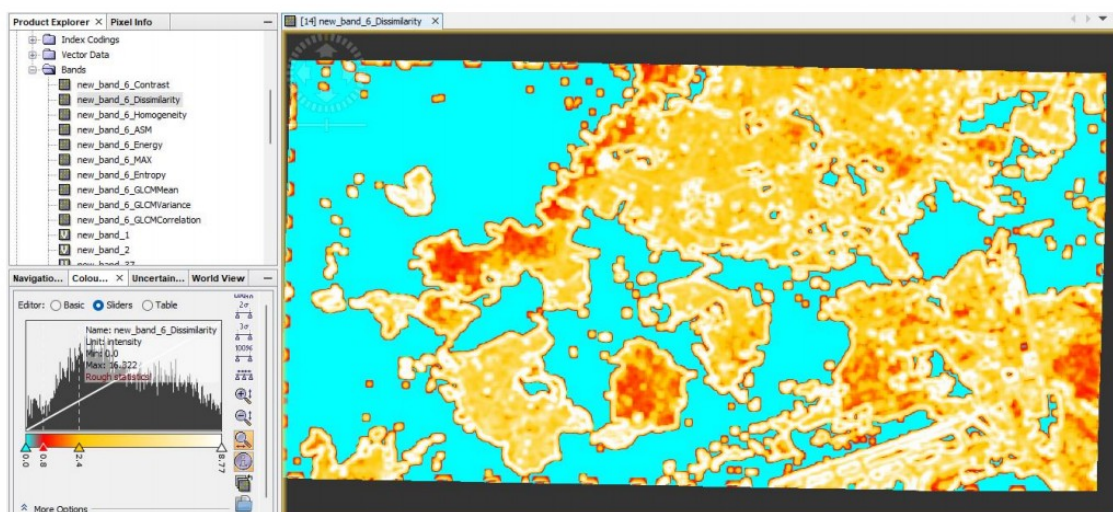


Figura 42: Resultados finales de la aplicacion de filtro GLCM de disimilitud y post-proceso (apertura) en la zona piloto.

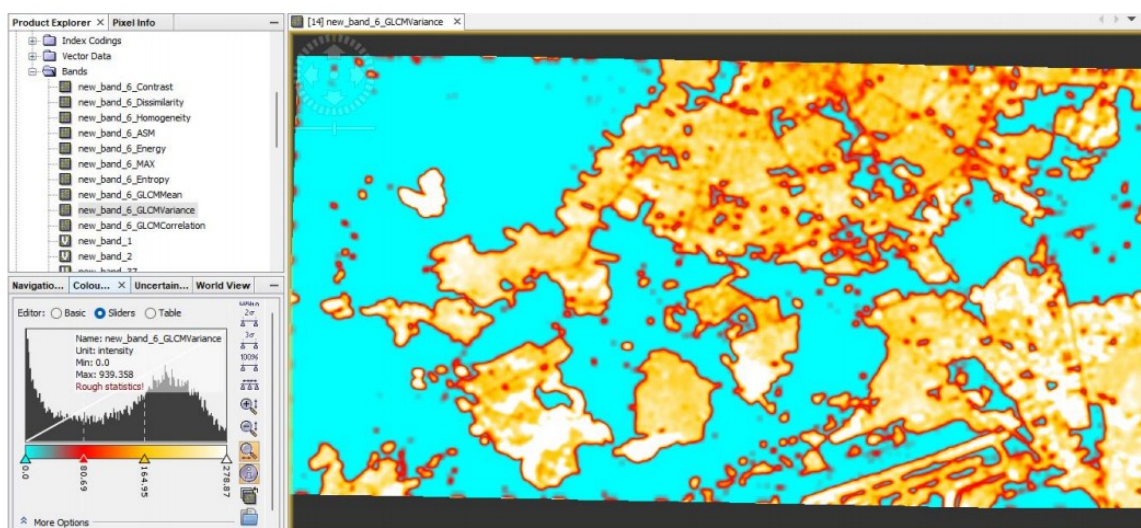


Figura 43: Resultados finales de la aplicacion de filtro GLCM de varianza y post-proceso (apertura) en la zona piloto

5.3.4 Estimación de la probabilidad de barrios de infravivienda (*slums*)

La probabilidad se estima normalizando los resultados de los filtros anteriores, entre un valor de 0 y 1. Para ello se determinan los valores máximos y mínimos de las imágenes filtradas, y se estima por consulta visual el valor de cada filtro que es más representativo de las zonas de *slum*. Sin embargo, para nuestro trabajo se elige el filtro de disimilitud, con una ventana de filtrado (15x15) que es la más adecuada y un post-procesado de apertura para evitar tener muchos residuos en nuestra imagen. Partiendo de esta imagen se aplica la siguiente expresión en Band Maths:

«if \$25.new_band_DILATE3_Dissimilarity < 1 then ((\$25.new_band_DILATE3_Dissimilarity - 0) / (1-0)) else (20 - \$25.new_band_DILATE3_Dissimilarity) / (20-1)»

Siendo 0 el valor mínimo; 20 el valor máximo; 1 el valor más probable para los *slums*.

El resultado se muestra en la Figura 44.

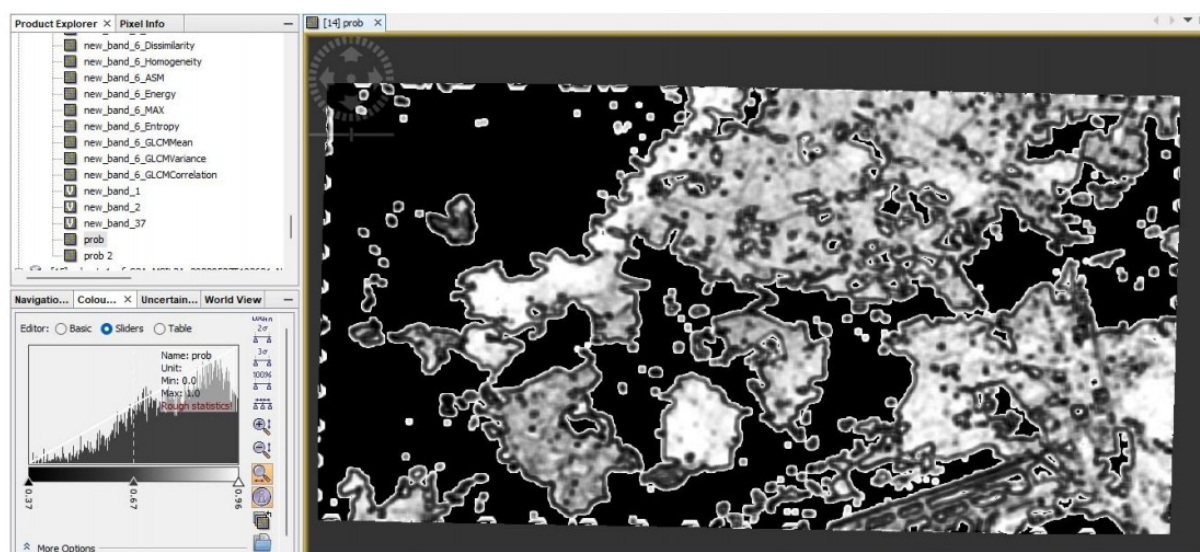


Figura 44: Imagen preliminar de la probabilidad de *slums*

No obstante, como aparecen residuos negativos y la escala lineal no es adecuada para definir la probabilidad se aplica una segunda expresión para determinar la probabilidad final:

«(if \$25.PROB < 0 then 0 else pow(10 , \$25.PROB) / 10)»

El resultado se muestra en la Figura 45.

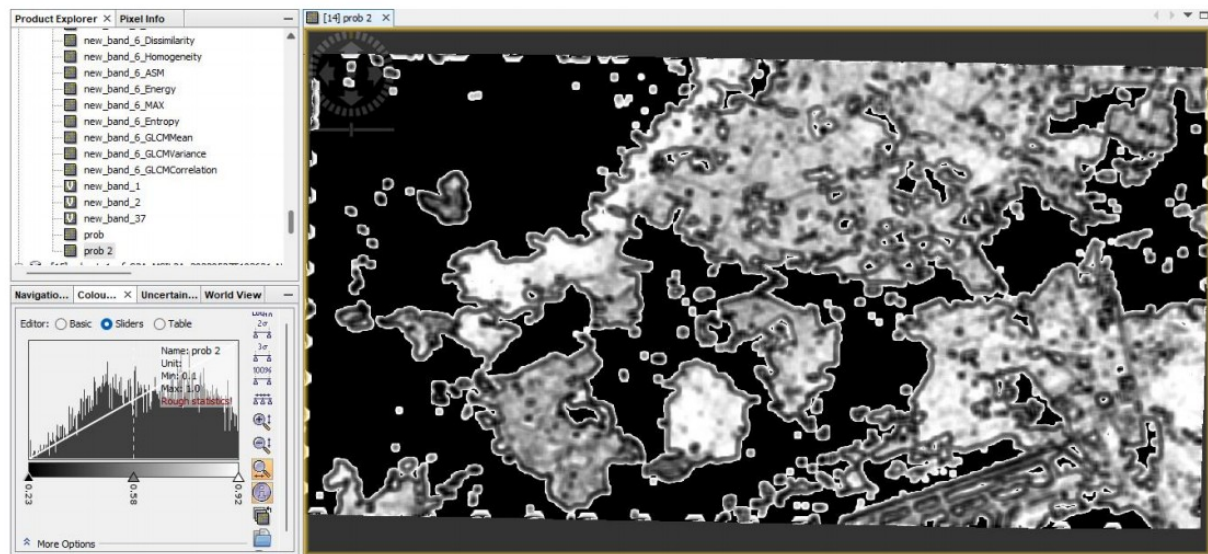


Figura 45: Imagen final de la probabilidad de slum

Los modelos resultantes en SNAP, se pueden exportar como imagen en formato TIFF, de forma que puedan ser visualizados desde QGIS, con una mejor presentación (Figura 46). Se emplea una escala de color bimodal de azul-amarillo-rojo (de más a menos probabilidad de encontrar slums)

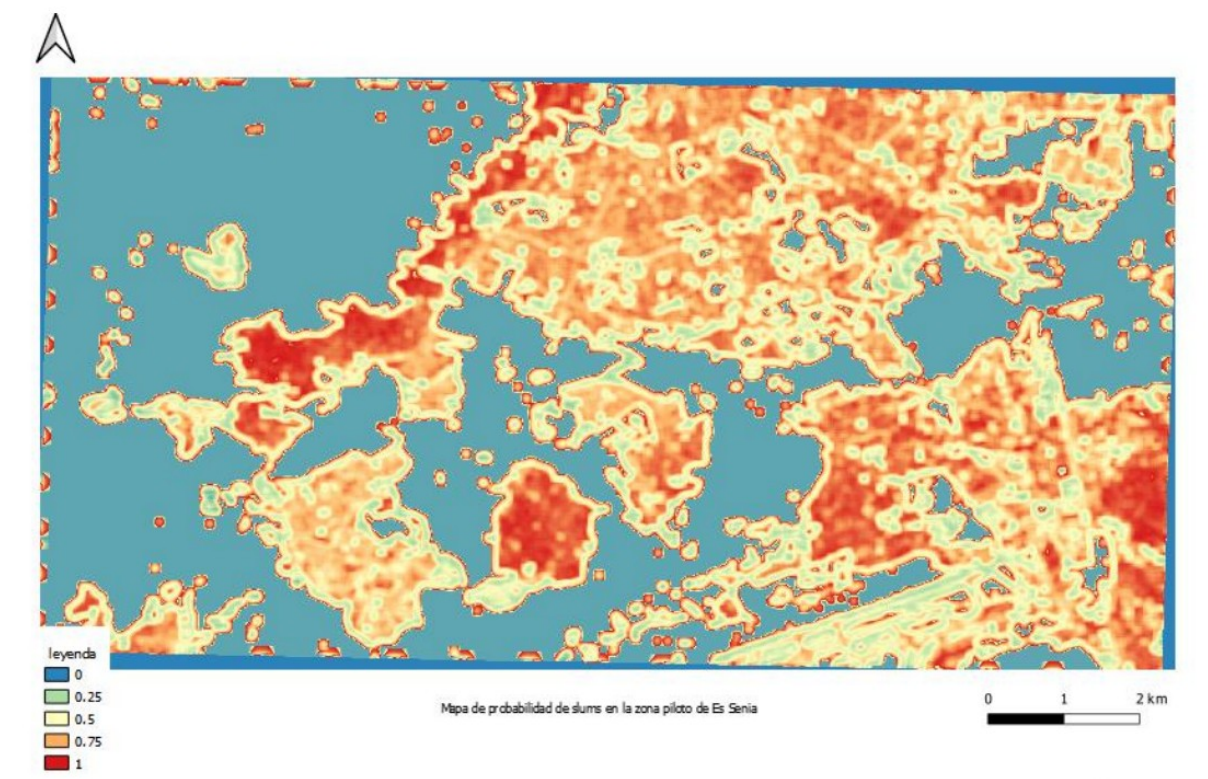


Figura 46: Mapa de probabilidad de slums en la zona piloto de Es senia

5.3.5 Validación de los resultados

La validación del mapa de probabilidad de encontrar slums, se realiza mediante la obtención de la matriz de confusión entre el mapa de probabilidad y la verdad terreno. Esta verdad terreno se obtiene mediante fotointerpretación a partir de las imágenes en color (real y falso) de Sentinel-2 y muestreo aleatorio de puntos dentro de estos polígonos y fuera de ellos.

5.3.5.1 Digitalización de polígonos y disolución

En primer lugar, se digitalizan algunos polígonos de áreas claramente identificadas como barrios de infravivienda (*slums*) en el área de estudios, sobre las imágenes de Sentinel-2, banda 8, aunque apoyándonos en las composiciones en color.

En segundo lugar, los polígonos se someten a la herramienta de disolución (*dissolve*) de QGIS, para que todos los polígonos queden integrados en una sola entidad (Figura 47).

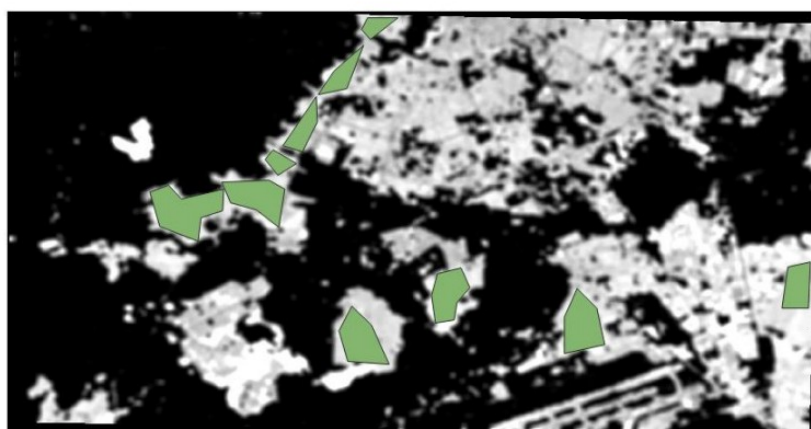


Figura 47. Polígonos digitalizados sobre la imagen de la banda 8.

De igual forma, se digitalizan polígonos externos a las zonas que se identifican como slum, tanto en zonas urbanas como no urbanas y se sometan al proceso de disolución.

5.3.5.2 Obtención de puntos aleatorios y extracción de probabilidad

Esta operación tiene como objetivo la creación de una tabla de valores en puntos de muestreo donde se disponga del valor de la probabilidad de slum y la verdad terreno.

Así, en primer lugar se crean dos capas nuevas de puntos con un número determinado de puntos de muestreo aleatorios dentro de los polígonos. Para ello se emplea la herramienta de QGIS “Puntos aleatorios dentro de polígonos” en la que se ha de fijar el número de puntos para cada entidad (hay solo una ya que los polígonos se han disuelto y la distancia mínima para evitar que los puntos estén demasiado cerca unos de otros (Figura 48). Se fijan 1000 puntos para cada una de las capas y no se fija distancia mínima.



Figura 48: Ventana de la herramienta de Puntos aleatorios dentro de polígonos

Se obtienen así dos capas de puntos, una para los polígonos consideradas como “slum” y otro para los polígonos “no slum”. A ambos, se les añade un campo de atributos, que en los “slum” toma un valor de 1 y en los “no slum” toma valor de 0. En la Figura 49 se muestra ambos conjuntos de puntos.

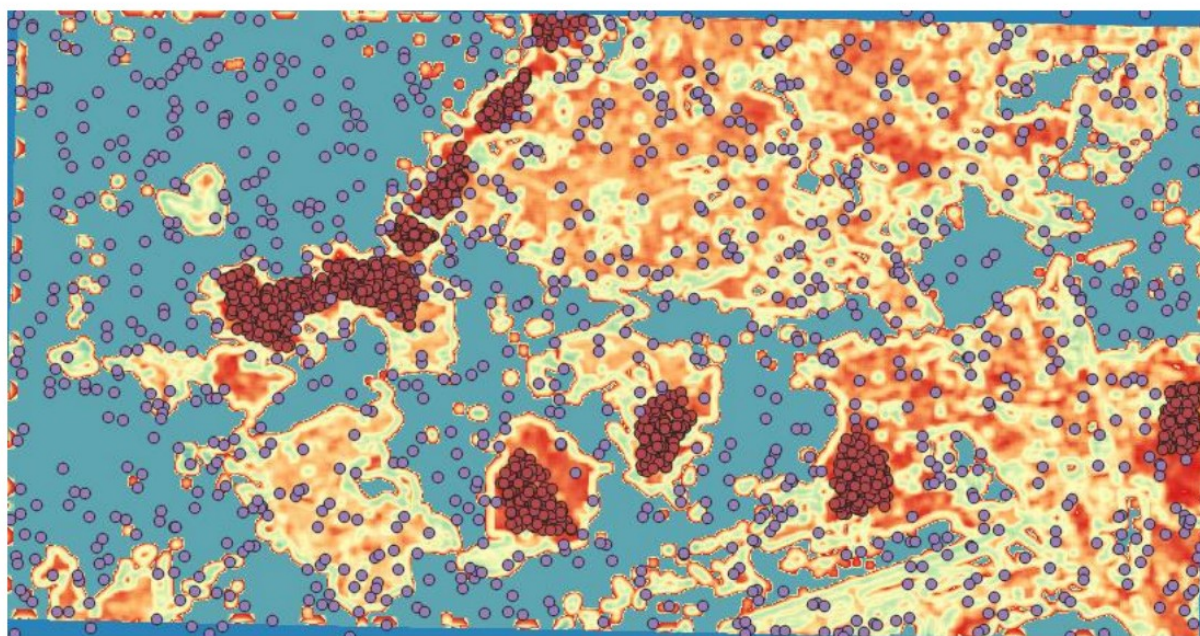


Figura 49. Capas de puntos aleatorio obtenidos para la validación en la zona piloto de Es Senia. En amarillo los puntos obtenidos como slum y en rojo como no slum.

A continuación, se procede a la extracción del valor de la susceptibilidad en cada punto mediante la herramienta Point Sampling Tool (plugin de QGIS), como se muestra en la Figura 50. Se ha de seleccionar como raster, el mapa de probabilidad. El resultado es que las tablas de los puntos tienen ahora el valor de la susceptibilidad (valor predicho) y el valor primario de slum/no slum (valor real).

Las dos tablas (dbf) se leen desde Excel, se fusionan y se preparan para introducirlas en el software R (Figura 51).

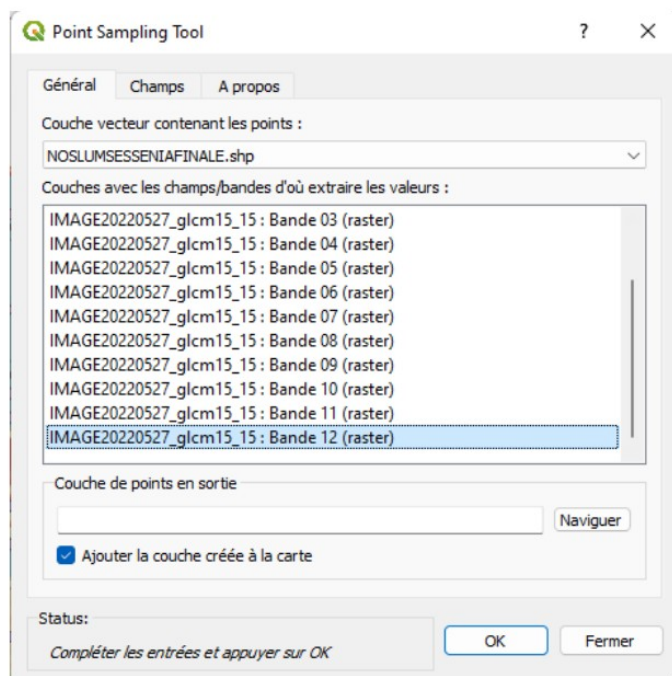


Figura 50: Ventana de Point Sampling Tool

	A	B			
1	PROB	vrai	1002	0,70828	0
2	0,85961	1	1003	0,64223	0
3	0,92499	1	1004	0,1	0
4	0,92696	1	1005	0,45933	0
5	0,93882	1	1006	0,73643	0
6	0,94681	1	1007	0,47962	0
7	0,8163	1	1008	0,99619	0
8	0,83555	1	1009	0,80463	0
9	0,66269	1	1010	0,76766	0
10	0,78742	1	1011	0,62717	0
11	0,89152	1	1012	0,5098	0
12	0,50144	1	1013	0,59507	0
13	0,95163	1	1014	0,45817	0
14	0,98653	1	1015	0,55795	0
15	0,78244	1	1016	0,1	0
16	0,89114	1	1017	0,1	0
17	0,88812	1	1018	0,44553	0
18	0,85743	1	1019	0,59155	0
19	0,88775	1	1020	0,1	0
20	0,9577	1	1021	0,1	0
			1022	0,87728	0
			1023	0,1	0

Figura 51. Tablas con los puntos de validación, el valor de probabilidad y la verdad terreno.

5.3.5.3 Obtención de la curva ROC y la matriz de confusión

La validación de los mapas de peligrosidad se realiza mediante la obtención de la curva ROC y el valor del área bajo la curva (AUC-ROC); y complementariamente con la matriz de confusión para un valor umbral de la probabilidad de 0,5 y los parámetros derivados de la matriz.



La matriz de confusión es una herramienta que permite medir la exactitud de los resultados obtenidos en un modelo predictivo, en nuestro caso el mapa de probabilidad de slum en relación con la realidad o verdad terreno. A partir de este análisis se obtienen los verdaderos positivos (slums que fueron predichos como tales), verdadero negativo (no slums predichas como tales), falso negativo (slums que no fueron predichos y se asignaron como no slums), falso positivo (zonas no slums que fueron predichas como zonas de slum). A partir de ella se pueden obtener una serie de parámetros derivados de la matriz como:

- La sensibilidad o recuerdo en la recuperación de información (recall) refleja la proporción de casos positivos que fueron correctamente identificadas por el algoritmo, es decir la razón entre verdaderos positivos (TP) y total de positivos reales (P):

$$\text{Sensibilidad (TPR)} = TP/P = TP/(TP+FN)$$

- La especificidad refleja la proporción de casos negativos catalogados correctamente, es decir la razón entre los verdaderos negativos (TN) respecto al total negativos (N) reales:

$$\text{Especificidad (TNR)} = TN/N = TN/(FP+TN)$$

- Relacionada con la especificidad se encuentra la tasa o razón de falsos positivos (FP):

$$FNR = FP/N = (N-TN)/(FP+TN) = 1 - \text{Especificidad (TNR)}$$

- La precisión, entendida como la dispersión del conjunto de valores obtenidos a partir de mediciones repetidas de una magnitud, cuanto menor es la dispersión mayor es la precisión. Dicho de otra forma, es la proporción de verdaderos positivos (TP) respecto al total de positivos predichos por el modelo (P'), por lo que también se conoce como poder predictivo positivo del modelo:

$$\text{Precisión (PPV)} = TP/P' = TP/(TP+FP)$$

- La exactitud que es la representación de la proporción de resultados verdaderos, tanto verdaderos positivos como verdaderos negativos dividido entre el número total de casos examinados:

$$\text{Exactitud} = (TP+TN)/(P+N) = (TP+TN)/(TP+FP+FN+TN)$$

- Finalmente, el puntaje (score) F1, se define como la media armónica de las métricas de precisión y sensibilidad:



$$F1 = 2 * (PPV * TPR) / (PPV + TPR) \text{ (ecuación 8)}$$

Se ha obtenido la matriz de confusión para un umbral del mapa de susceptibilidad de 0,5, lo que significa que se considera slum aquellas zonas con probabilidad superior a 0,5; y como no slum, las zonas con probabilidad inferior a 0,5. Como este valor es un tanto arbitrario, hay otra posibilidad que es ensayar distintos umbrales, evaluando su rendimiento a través de la curva ROC y el AUC, que ahora se describe.

La curva ROC (receiver operating characteristic curve) es un gráfico que muestra el rendimiento de un modelo de clasificación en todos los umbrales de clasificación. Esta curva traza dos parámetros:

- La tasa de verdaderos positivos
- La tasa de falsos positivos

La tasa de verdaderos positivos (TPR) es la siguiente:

$$TPR = TP / (TP + FN)$$

La tasa de falsos positivos (FPR) se define como sigue:

$$FPR = FP / (FP + TN)$$

Siendo TP los verdaderos (true) positivos; FN, los falsos negativos; FP, los falsos positivos; y TN, los verdaderos negativos.

Una curva ROC representa la TPR en función de la FPR en diferentes umbrales de clasificación. Si se reduce el umbral de clasificación, se clasifican más elementos como positivos, aumentando tanto los falsos como los verdaderos positivos.

El valor AUC significa "Área bajo la curva ROC" (Area Under the ROC-Curve), y mide toda el área bidimensional bajo la curva ROC completa (de acuerdo al cálculo integral) desde (0,0) hasta (1,1).

El AUC proporciona una medida agregada de rendimiento para todos los posibles umbrales de clasificación. El valor del AUC varía de 0 a 1. Un modelo con un 100% de predicciones falsas tiene un CUA de 0,0; un modelo con un 100% de predicciones correctas tiene un CUA de 1,0.

El AUC es deseable por las dos razones siguientes:

- El AUC es invariable en cuanto a la escala. Mide la calidad de la clasificación de las predicciones, más que su valor absoluto.



- El AUC es invariable con respecto al umbral de clasificación. Mide la calidad de las predicciones del modelo, independientemente del umbral de clasificación elegido.

Tanto la matriz de confusión y sus parámetros, como la curva ROC y el valor del área bajo la curva (AUC) se calculan mediante un sencillo script de R, a partir de la tabla Excel importada de este software (Figura 52). La salida del resultado de la curva ROC y el valor del AUC se muestra en la Figura 53.

```
> View(PROJETESSENIA2)
> pred <- prediction(PROJETESSENIA2$PROB, PROJETESSENIA2$vrai)
> perf <- performance(pred, "tpr", "fpr")
> plot(perf,
+ main = "Curva ROC ",
+ xlab="Tasa de falsos positivos (1 - especificidad)",
+ ylab="Tasa de verdaderos positivos (sensibilidad)")
> abline(a=0,b=1,col="blue",lty=2)
> grid()
> auc <- as.numeric(performance(pred,"auc")@y.values)
> legend("bottomright",legend=paste(" AUC =",round(auc,4)))
> tabla <- table(REAL=PROJETESSENIA2$vrai,PREDICHO=PROJETESSENIA2$PROB>0.5)
> performance <- data.frame(tabla)
> TN <- sum(performance$Freq[1])
> FN <- sum(performance$Freq[2])
> FP <- sum(performance$Freq[3])
> TP <- sum(performance$Freq[4])
> precision = TP/(TP+FP)
> recall = TP/(TP+FN)
> accuracy = (TP+TN)/(TP+TN+FP+FN)
> f1_score = (2*precision*recall)/(precision+recall)
> tabla
  PREDICHO
REAL FALSE TRUE
0      559 441
1      18 982
> precision
[1] 0.6900914
> recall
[1] 0.982
> accuracy
[1] 0.7705
> f1_score
```

Figura 52: Script de R para la obtención de la matriz de confusión y la curva ROC

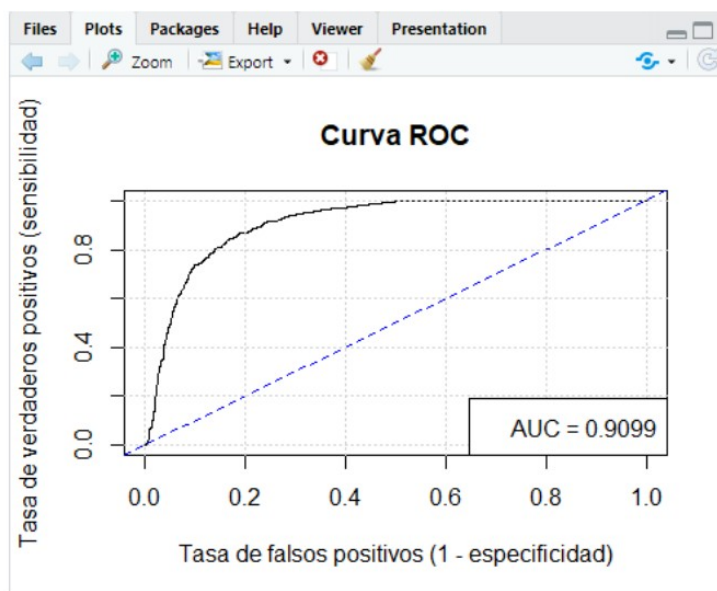


Figura 53: Curva ROC y valor del área bajo la curva (AUC) en la zona piloto

6 Resultados y discusión

6.1 Mapas de probabilidad a los slums

6.1.1 Zona piloto (Es Senia)

La Figura 54 muestra los resultados obtenidos utilizando diferentes operadores de filtrado: contraste, disimilitud, entropía y varianza en la zona piloto (Es Senia). Se realizaron en imágenes preprocesadas con los filtros GLCM y postprocesadas con el filtro de apertura mediante la ventana de filtro 15*15.

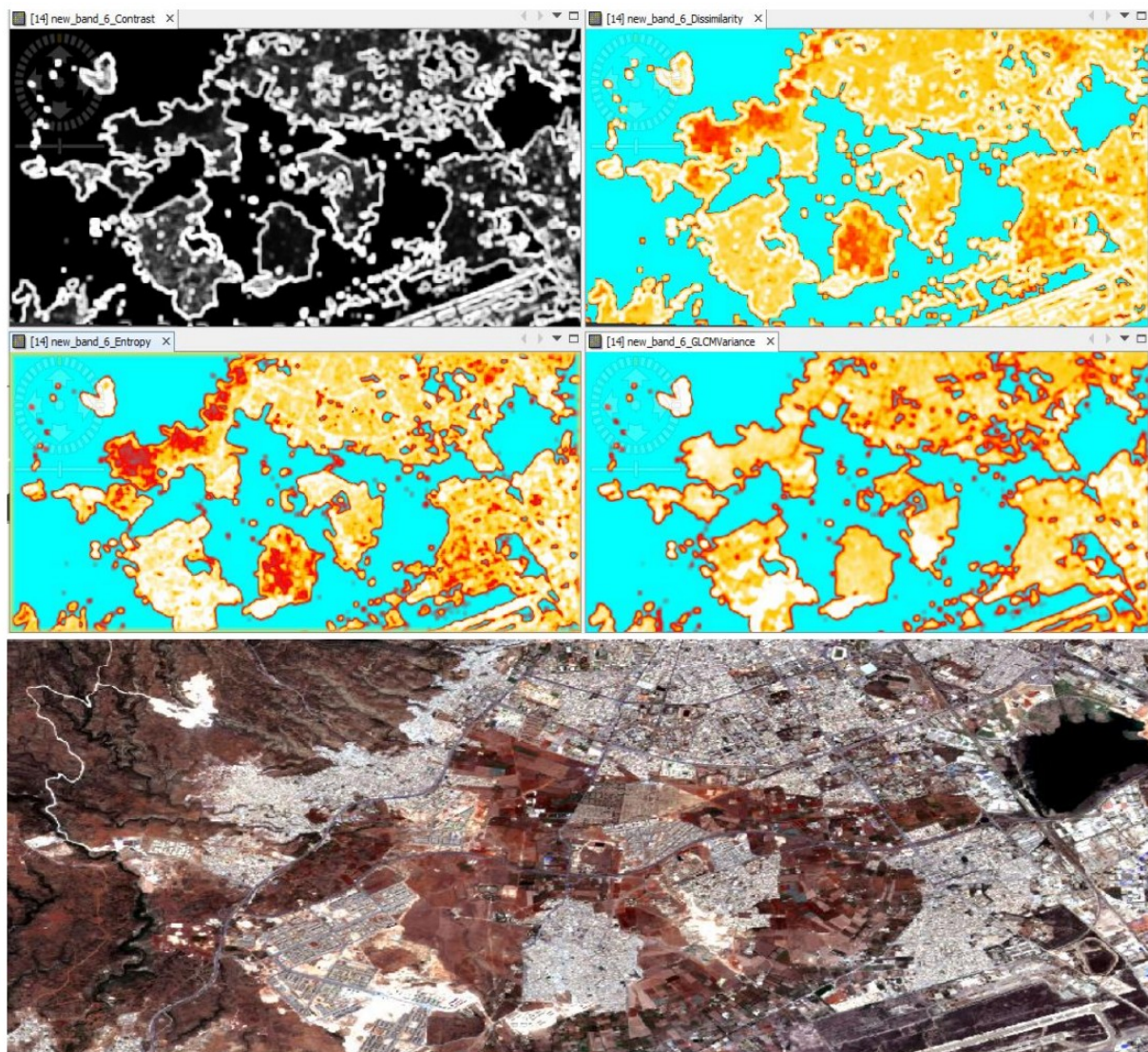


Figura 54: Filtros texturales en la zona piloto mediante distintos operadores e imagen en color real (abajo)

Los operadores que proporcionan mejores resultados visualmente son la entropía y la disimilitud. En ellos, se identifican claramente los slums en el noroeste y en el sur de la zona. Sin embargo, la disimilitud parece discriminar algo mejor, ya que las zonas con slum aparecen mas claramente

identificadas (tonos más rojos) y las zonas de nueva vivienda (barrios nuevos en el suroeste y en el sureste) aparece como zonas menos probables de slum (tonos amarillos).

En la figura 55 se muestra el mapa de probabilidad mediante el método de disimilitud, sometido a un post-proceso que reduce algo los bordes más disímiles o contrastados de las zonas urbanas y elimina objetos de tamaño pequeño.

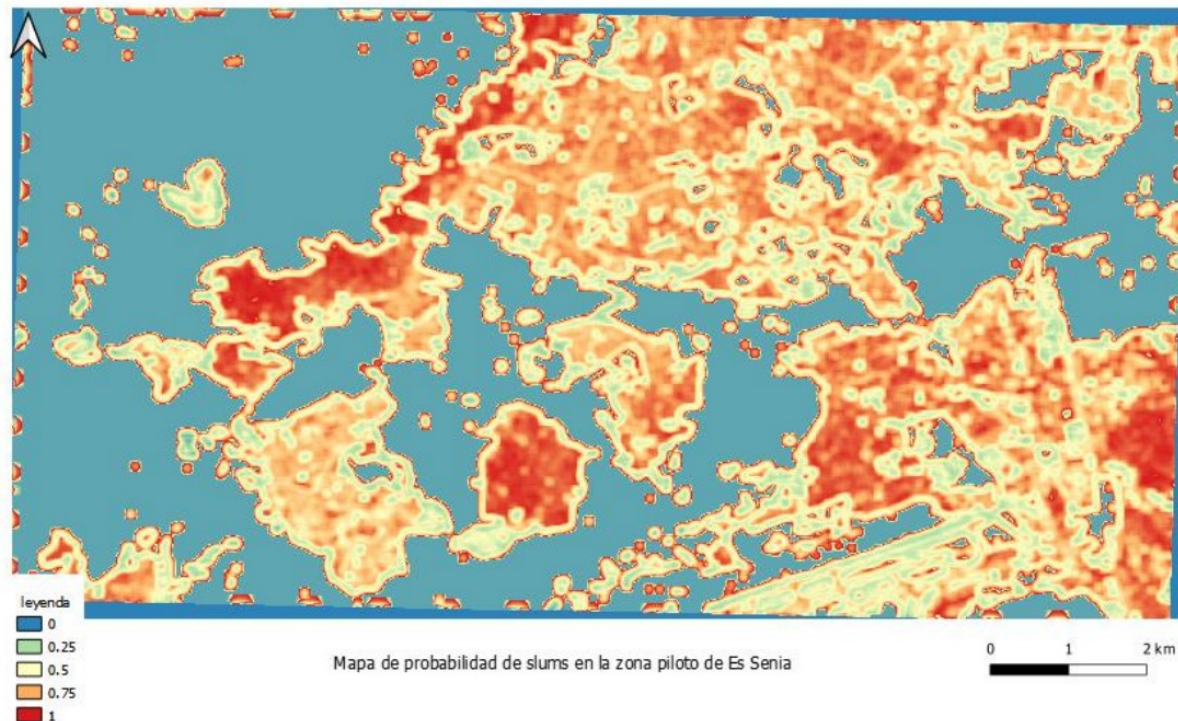


Figura 55: Mapa de peligrosidad de slums en la zona piloto de Es Senia

6.1.2 Zona de Orán

En la figura 56 se muestran los resultados obtenidos mediante diversos operadores de filtrado: contraste, disimilitud, entropía y varianza en la zona completa de Orán. Se han hecho sobre imágenes pre-procesadas mediante filtros glcm, y postprocesadas con el filtro de apertura mediante la ventana de filtro 15*15.

Se trata de una imagen más amplia que la anterior, en el que además de las zonas de slum del noroeste, y del sur (ahora más en el suroeste) aparecen otras al este de la ciudad y al sur de la misma (al este de las anteriores). El centro de ciudad de Orán queda libre de slums, así como otros barrios periféricos de nuevo construcción en el sur.

Por su parte, en la figura 57 se muestra el mapa de probabilidad de toda la zona, habiendo seleccionado el método de la disimilitud por iguales razones que antes (es la misma zona ampliada).

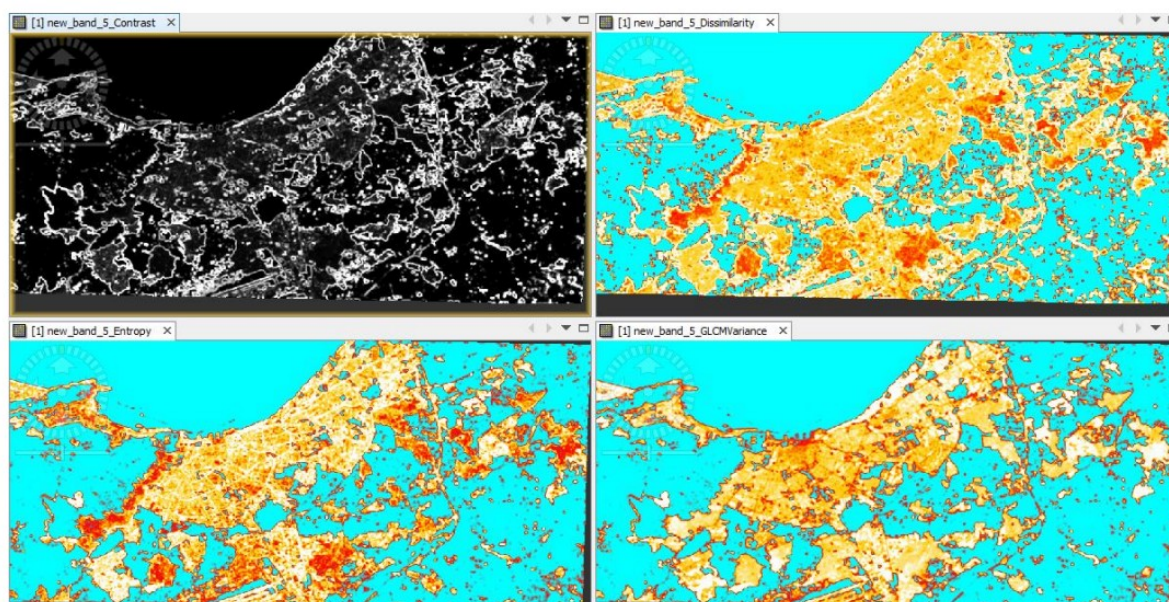


Figura 56: Filtros texturales en la zona de Orán mediante distintos operadores

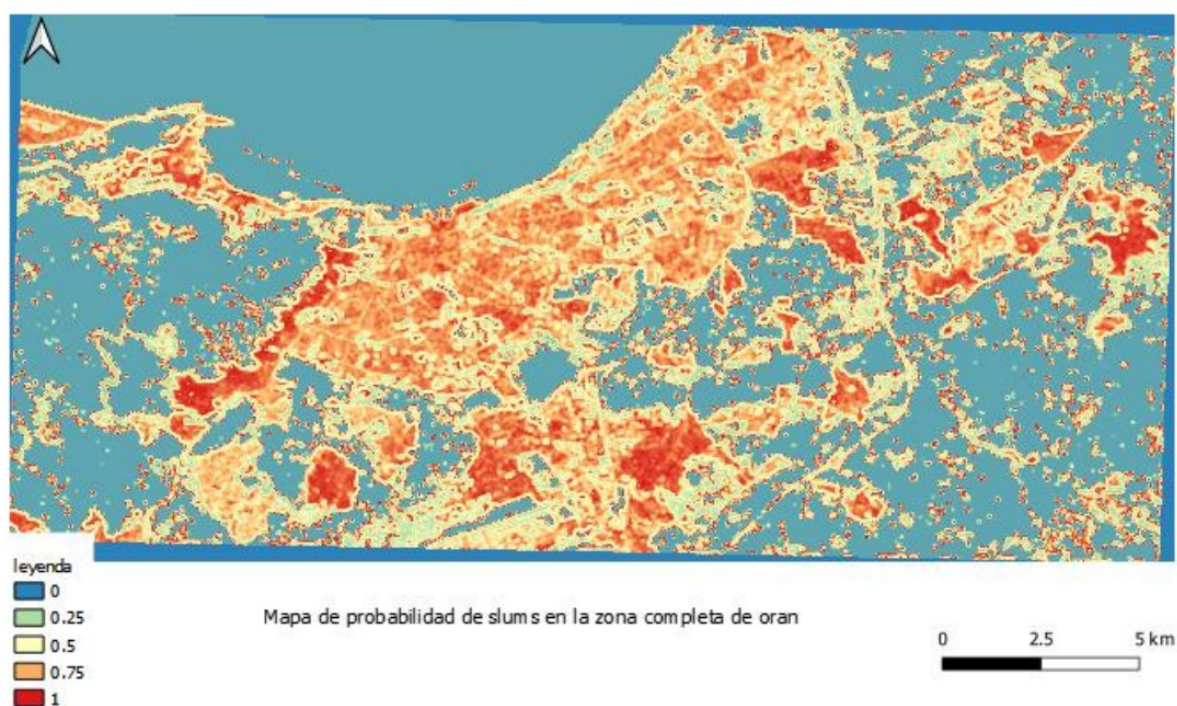


Figura 57: Mapa de peligrosidad de slums en la zona completa de Oran

6.1.3 Zona de Argel

En la figura 58 se muestran los resultados obtenidos mediante diversos operadores de filtrado: contraste, disimilitud, entropía y varianza en la zona de argel. Se han hecho sobre imágenes pre-procesadas mediante filtros glcm, y post-procesadas con con el filtro de apertura mediante la ventana de filtro 15*15.

Como en las zonas anteriores, los operadores que proporcionan mejores resultados visualmente son la entropía y la disimilitud, especialmente la segunda en la que se identifican los slums de la periferia de la ciudad al noreste, al noroeste y al sur, dejando más libre la zona centro de la ciudad.

Por su parte, en la figura 59 se muestra el mapa de probabilidad de toda la zona, habiendo seleccionado el método de la disimilitud que proporciona una mejor discriminación.

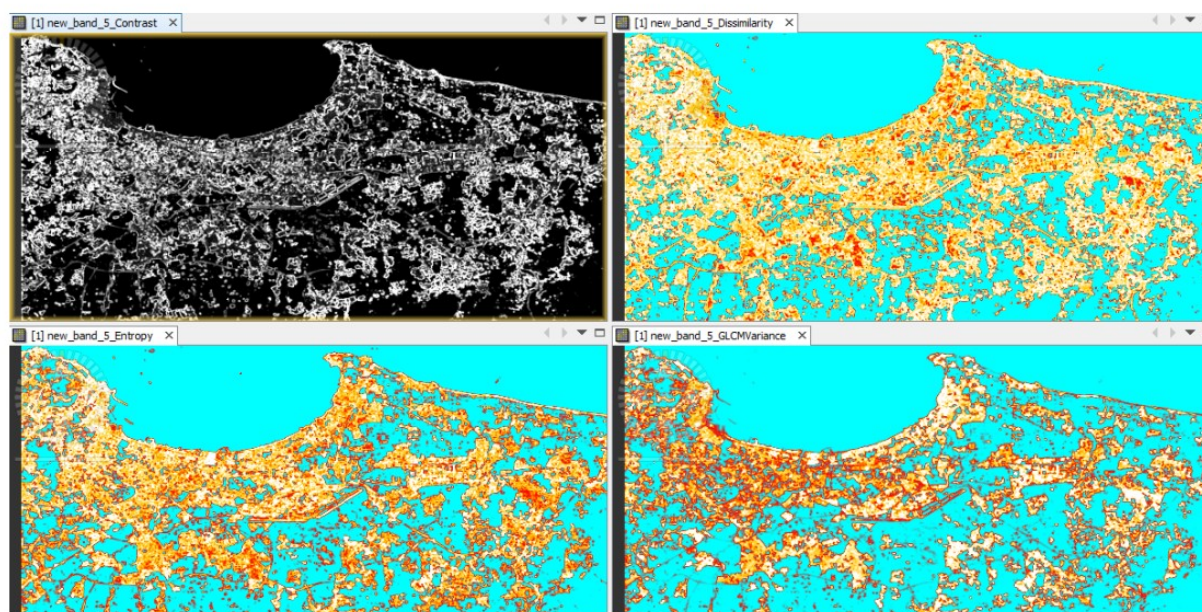


Figura 58: Filtros texturales en la zona de Argel mediante distintos operadores

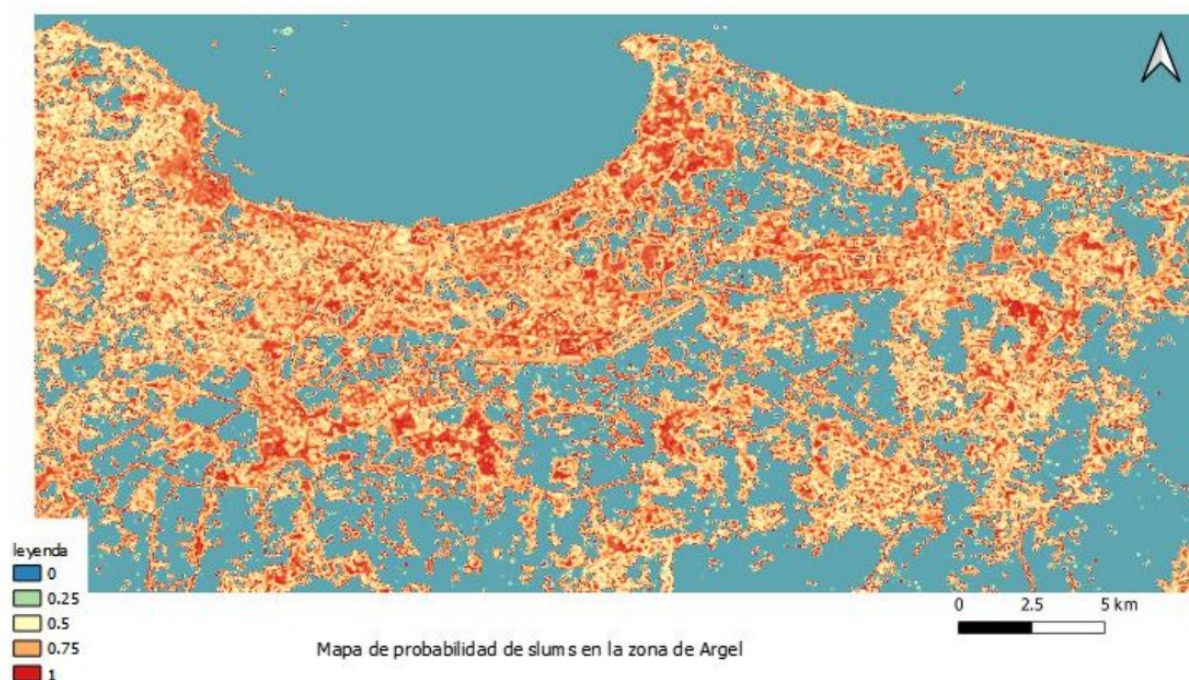


Figura 59: mapa de peligrosidad de slums en la zona de Argel

6.2 Validación

6.2.1 Zona piloto (Es Senia)

En la figura 60 se presenta la localización de los puntos de muestreo en la zona piloto. Por su parte, en la tabla 5 se muestran la matriz de confusión y en la figura 61 la curva AUC ROC obtenida.

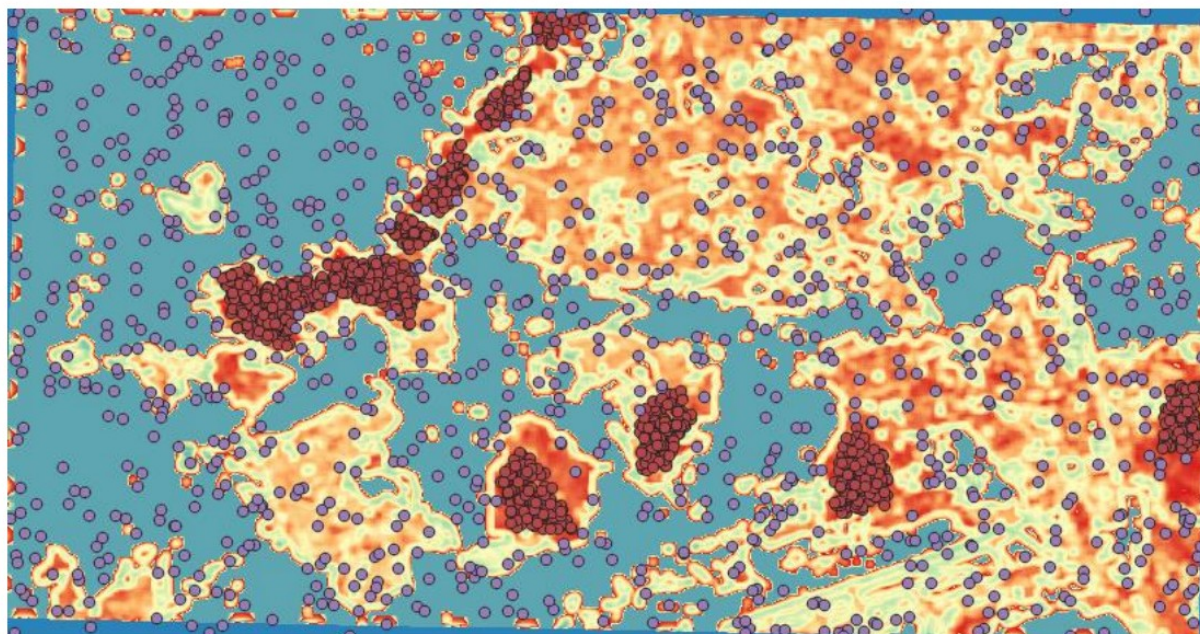


Figura 60: Puntos de muestreo para validación en la zona piloto

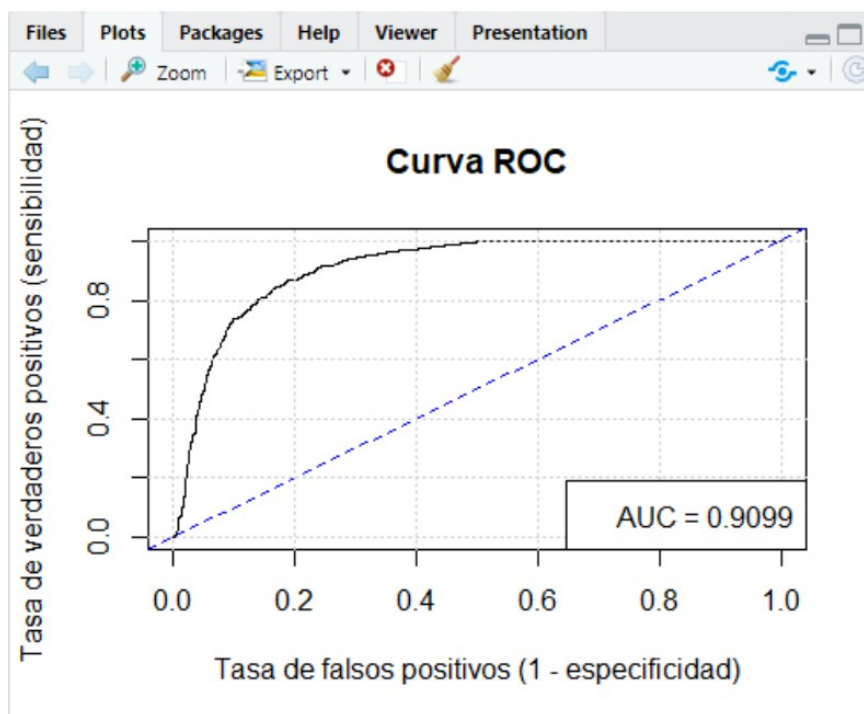


Figura 61: Curva ROC y resultado de la AUC en la zona piloto

	Predicción		Precisión	Sensibilidad	Exactitud	F1 Score
Realidad	Positivos	Negativos	0,69	0,98	0,77	0,81
Positivos	982	441				
Negativos	18	559				

Tabla 5: Curva ROC y resultado de la AUC en la zona piloto

Los resultados son bastante buenos, ya que el valor del AUC ROC es bastante elevado (0,9099). Hay que tener en cuenta que este parámetro es el más global, ya que los parámetros restantes se refieren a una partición del modelo por un único umbral de 0,5. No obstante la exactitud y el F1 presentan igualmente valores aceptables (0,77 y 0,81). La precisión es de 0,69 y la sensibilidad de 0,98. Según estos valores podemos considerar que el método desarrollado da unos muy buenos resultados en la zona piloto, lo que nos induce a utilizarlo en las otras zonas.

6.2.2 Zona de Orán

En la figura 62 se presenta la localización de los puntos de muestreo en la zona de Orán y en la figura 63 la curva AUC ROC obtenida.

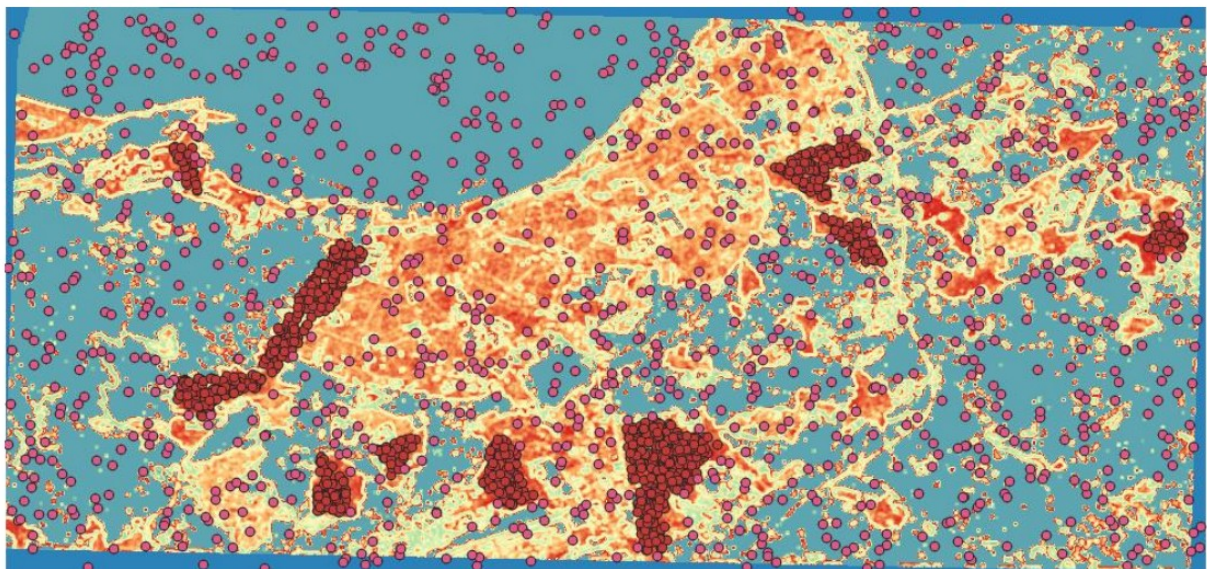


Figura 62: Puntos de muestreo para validación en la zona de Oran

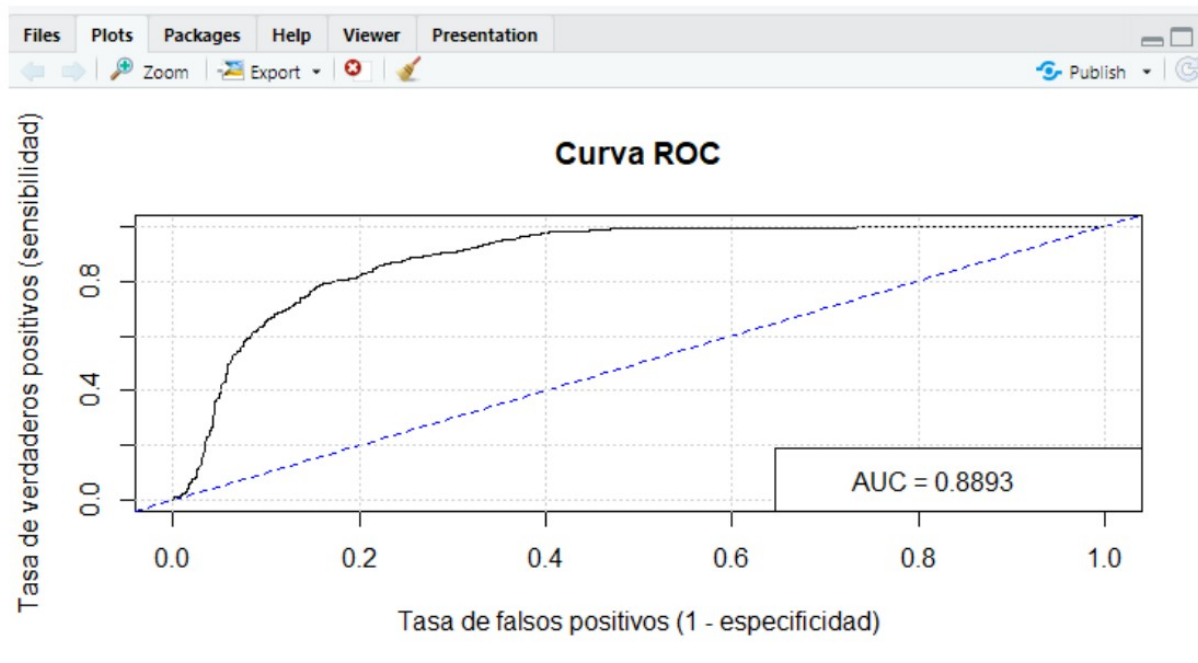


Figura 63: Curva ROC y resultado de la AUC en la zona de Orán

En este caso, la zona ampliada de Orán, el resultado de la curva ROC es también muy bueno de 0,8893 para el área bajo la curva, lo que permite validar el mapa de probabilidad de slums en esta zona.

6.2.3 Zona de Argel

En la figura 64 se presenta la localización de los puntos de muestreo en la zona de Argel y en la figura 65 la curva AUC ROC obtenida.

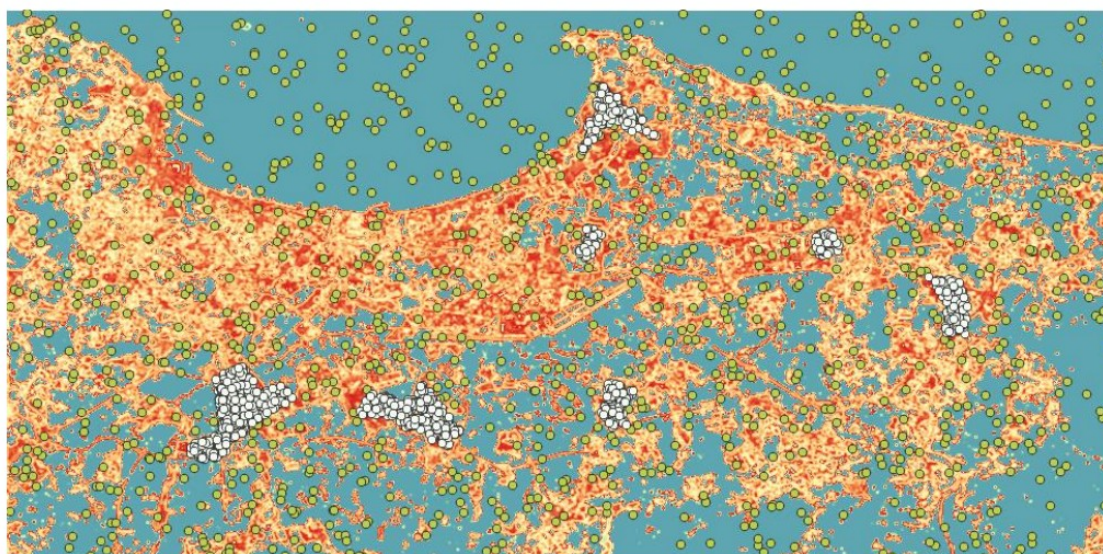


Figura 64: Puntos de muestreo para validación en la zona de Argel

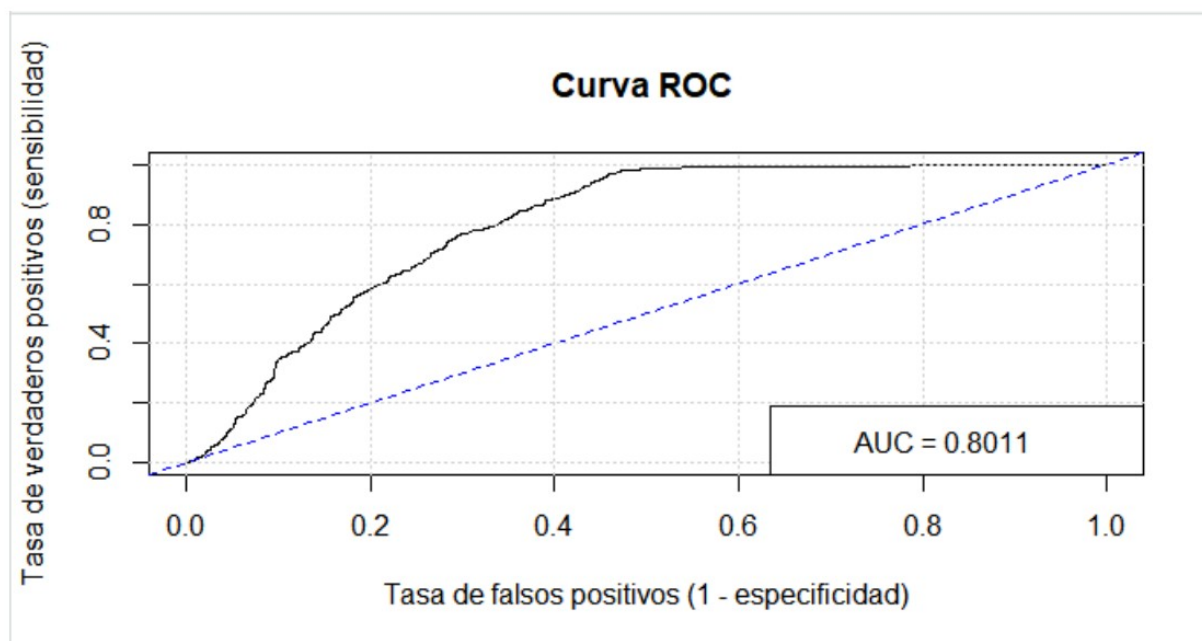


Figura 65: Curva ROC y resultado de la AUC en la zona Argel

Al igual que en las zonas anteriores el valor del area bajo la curva es bastante elevado (0,8011), aunque algo más bajo, lo que no impide validar el mapa de probailidad a los slusm de la zona de Argel.



7 Conclusiones

En este trabajo se ha planteado la identificación de barrios marginales y precarios en las ciudades, concretamente los barrios de infravivienda (*slums*), un tema de gran importancia a nivel mundial y en Argelia en particular. Así, en 2010, la cifra de infraviviendas para todo el país era de 554.000, incluyendo todos los tipos, con un 60% situadas en zonas urbanas y un 40% en zonas rurales. Es por lo tanto un problema en el que es necesario actuar, para mejorar la calidad de vida de un gran número de personas.

Uno de los primeros pasos es la identificación, cartografía y cuantificación del fenómeno, lo que se plantea en este trabajo mediante técnicas semiautomáticas de teledetección y análisis de imágenes. De acuerdo a la bibliografía consultada, en la que aparecen ejemplos de teledetección con imágenes de alta y media resolución, en este trabajo se propone la utilización de imágenes de Sentinel-2 con una adecuada resolución espacial (10-20 m) para la detección de barrios de infravivienda (*slums*), no tanto para viviendas específicas; así como una excelente resolución espectral y temporal (5 días) y una disponibilidad total de las imágenes en servidores de datos de forma gratuita. Se plantea también la utilización de imágenes de radar de Sentinel-1 con igualmente una buena resolución espacial y temporal (6 días).

Además de las imágenes de Sentinel, se ha apostado por la utilización de software libre, en primer lugar, de SNAP, paquete desarrollado por la ESA y especialmente indicado para el procesamiento y análisis de estas imágenes; QGIS, un software SIG también de carácter libre, que se emplea para la presentación cartográfica de los resultados; y R, un paquete estadístico igualmente libre, que se emplea para la parte final de la validación de los resultados.

Con estos medios, se ha desarrollado una metodología que comprende una primera fase de preparación y pre-procesamiento de las imágenes, más sencillo en las imágenes ópticas de Sentinel-2 y más compleja en las imágenes de radar de Sentinel-1, con vistas a preparar las imágenes para su clasificación. Una vez preparadas las imágenes se han realizado clasificaciones supervisadas, empleando un método contrastado como los bosques aleatorios (*random forest*), tanto de unas como de otras. La clasificación de imágenes de Sentinel-2 resultan válidas para detectar zonas urbanas frente a otras coberturas del suelo, pero no para diferenciar entre distintos tipos de zonas urbanas, como los *slum*. Las imágenes de Sentinel-2 permiten también identificar zonas urbanas en su conjunto, pero con mucho más ruido, por lo que se descarta su uso.



La detección adecuada de zonas urbanas con Sentinel-2 permite la realización de máscaras para centrar el análisis en dichas zonas descartando las demás, lo que evita ruido en los procesos siguientes. En estas zonas se propone la aplicación de filtros texturales de co-ocurrencia de niveles de gris (GLCM) para la diferenciación de los slum, combinado con operaciones de morfología matemática, tanto como pre-proceso, como post-proceso.

Los mejores resultados se obtienen en los filtros de entropía y de disimilitud, especialmente en este último, combinado con operadores de apertura y cierre (morfología matemática). Las imágenes de disimilitud se transforman en mapas de probabilidad de slums mediante operaciones de reclasificación lógica y re-escalado, que muestran en cada pixel un valor de probabilidad ente 0 y 1.

Los mapas de la zona piloto se someten a un proceso de validación mediante muestreo aleatorio de puntos, que permite determinar la matriz de confusión y sus parámetros (precisión, sensibilidad, exactitud y F1-score), así como la curva ROC y el área bajo la curva (AUC). Este último parámetro (AUC-ROC), que es una medida más global de la validación, proporciona valores muy elevados (por encima de 0,80-0,90), que permiten validar los mapas de probabilidad de slum, por lo tanto, con un 80-90% de fiabilidad.

Esta validación se ha hecho no solo en el área piloto (Es Senia), sino en las áreas urbanas de las dos ciudades más importantes de Argelia que son la capital Argel y la ciudad de Orán, muy afectada por el fenómeno de la infravivienda.

Como conclusión final, se ha desarrollado y validado una metodología para la detección de barrios de infravivienda, mediante un procedimiento sencillo y aplicable a la resolución de un problema acuciante y de gran importancia a nivel global y, especialmente, es países en desarrollo como Argelia. Es un primer paso para la identificación del problema y a partir de ahí, la actuación por las autoridades competentes.

La metodología es además mejorable en el futuro, con la utilización de imágenes de mayor resolución, el ensayo de nuevos procesos de morfología matemática y análisis textural; o, sobre todo, con la aplicación de técnicas de aprendizaje automático profundo como las redes neuronales convolucionales, muy apropiadas para la identificación de patrones complejos en imágenes.



8 Bibliografía

BEO (2022). SLUMAP : Cartographier les quartiers défavorisés pour définir des politiques de lutte contre la pauvreté. Belgian Earth Observation. <https://eo.belspo.be/fr/actualites/slumap-cartographier-les-quartiers-defavorises-pour-definir-des-politiques-de-lutte> (Último acceso dic. 2022).

[Breiman, L. Random Forests. *Machine Learning* 45, 5–32 \(2001\)](#)

Braun A and Offermann E (2022). Polarimetric information content of Sentinel-1 for land cover mapping: An experimental case study using quad-pol data synthesized from complementary repeat-pass acquisitions. *Front. Remote Sens.* 3:905713. doi: 10.3389/frsen.2022.905713

DEA (2022). Urban area mapping using Sentinel 1 data. Digital Earth Africa. https://docs.digitalearthafrika.org/en/latest/sandbox/notebooks/Real_world_examples/Radar_urban_area_mapping.html (Último acceso dic. 2022)

Engstrom, R., Sandborn, A., Qin Yu, Burgdorfer, J. (2015). Mapping slums using spatial features in Accra, Ghana. *Joint Urban Remote Sensing Event (JURSE)*, 2015, pp. 1-4, doi: 10.1109/JURSE.2015.7120494.

Haralick, R.M.; Shanmugam, K. Dinstein, I. (1973). Textural Features for Image Classification. *IEEE Trans. on Systems, Man and Cybernetics*, Vol 3 , No. 6, pp. 610-621.

Kohli, D. (2015). *Identifying and classifying slums areas using remote sensing*. ITC dissertation No. 279 University of Twente (The Netherlands). 131 p.

Kuc, G. y Chormański, J. (2019). SENTINEL-2 imagery for mapping and monitoring imperviousness in urban areas, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLII-1/W2, 43–47, <https://doi.org/10.5194/isprs-archives-XLII-1-W2-43-2019>, 2019.

Najmi, A. (2021). *Integrating remote sensing and street view images to map slums using deep learning approach*. Master's Dissertation University of Twente (The Netherlands). 75 p.

Reiu, A. (2017). *Classification for urban areas from Sentinel-1 coherence maps*. Bachelor's thesis University of Tartu (Estonia). 32 p.



SLUMAP (2022). SLUMAP Project: Remote Sensing for Slum Mapping and Characterization in sub-Saharan African Cities. <https://slumap.ulb.be/> (Último acceso dic. 2022)

Veille Carto 2.0 (2022). L'intégration cartographique des bidonvilles. <https://veillecarto2-0.fr/2022/01/14/integration-cartographique-des-bidonvilles/> (Último acceso dic. 2022)