

Clasificación de usos del suelo a partir de imágenes Sentinel-2

Borràs, J.^{*1}, Delegido, J.¹, Pezzola, A.², Pereira, M.¹, Morassi, G.¹, Camps-Valls, G.¹

¹ Laboratorio de Procesado de Imágenes, Universidad de Valencia. C/ Catedrático José Beltrán, 2. 46980 Paterna, València - España.

² Instituto Nacional de Tecnología Agropecuaria. Estación Experimental Hilario Ascasubi. Laboratorio de Teledetección y SIG. Ruta Nac. 3 sur km. 794. (8142) Hilario Ascasubi, Argentina.

Resumen: Sentinel-2 (S2) es un nuevo satélite de la ESA que cuenta con 13 bandas proporcionando imágenes de alta calidad radiométrica y excelente resolución espacial (10 y 20 m) ideal para trabajos de clasificación. En este trabajo se han abordado dos objetivos: determinar el mejor método de clasificación con S2, y cuantificar su mejora respecto a otras misiones operativas, como SPOT. Para ello se han seleccionado cuatro clasificadores (LDA, RF, Árboles de decisión, K-NN) que se han aplicado en dos zonas agrarias: una en la huerta de Valencia (España) y otra en la región de Buenos Aires (Argentina). Se han probado todos los clasificadores usando, por una parte, todas las bandas de S2, y por otra usando sólo las cuatro que coinciden con SPOT. En todos los casos se han aplicado porcentajes entre el 10 y el 50% de datos de entrenamiento y usado el resto de datos como validación. Como resultado se ha generado un mapa de usos del suelo a partir del mejor clasificador, basándose en el índice Kappa, proporcionando información científicamente relevante como es el área ocupada por cada una de las clases.

Palabras clave: clasificación, teledetección, usos del suelo, índice Kappa, Sentinel-2.

Land use classification from Sentinel-2 imagery

Abstract: Sentinel-2 (S2), a new ESA satellite for Earth observation, accounts with 13 bands which provide high-quality radiometric images with an excellent spatial resolution (10 and 20 m) ideal for classification purposes. In this paper, two objectives have been addressed: to determine the best classification method for S2, and to quantify its improvement with respect to the SPOT operational mission. To do so, four classifiers (LDA, RF, Decision Trees, K-NN) have been selected and applied to two different agricultural areas located in Valencia (Spain) and Buenos Aires (Argentina). All classifiers were tested using, on the one hand, all the S2 bands and, on the other hand, only selecting those bands from S2 closer to the four bands from SPOT. In all the cases, between 10%-50% of samples were used to train the classifier while remaining the rest for validation. As a result, a land use map was generated from the best classifier, according to the Kappa index, providing scientifically relevant information such as the area of each land use class.

Keywords: classification, remote sensing, land use, Kappa Index, Sentinel-2.

1. Introducción

El uso del suelo es un criterio socioeconómico que hace referencia a la actividad dominante que

se da en un determinado lugar, e incluye usos como residencial, comercial, industrial, agrícola, etc., pudiendo estar estas categorías divididas en niveles más específicos. La cobertura del suelo es

To cite this article: Borràs, J., Delegido, J., Pezzola, A., Pereira, M., Morassi, G., Camps-Valls, G. 2017. Land use classification from Sentinel-2 imagery. *Revista de Teledetección*, 48, 55-66. <https://doi.org/10.4995/raet.2017.7133>

* Autor para la correspondencia: Jubohar@alumni.uv.es

un indicador biofísico que describe los materiales que cubren el territorio, como por ejemplo, asfalto, vegetación, suelo desnudo, agua, etc.

Según Comber *et al.* (2005), la cobertura del suelo es entendida de forma diferente según las disciplinas científicas. Los ecólogos definen la cobertura del suelo según la presencia y abundancia de especies vegetales, mientras que para los edafólogos es un indicador de la composición y los tipos de suelos. Por otra parte, en la bibliografía en castellano existe el término “ocupación del suelo” que según López González *et al.* (2002) hace referencia a la cobertura, entendida como los materiales que cubren el territorio, mientras que otros autores la consideran como una combinación de los conceptos de uso y cobertura (Del Bosque *et al.*, 2005). Es muy habitual la utilización conjunta de los términos uso y cobertura, ya que, en función del objetivo final de la cartografía, se definen leyendas donde coexisten clases que corresponden a usos con otras que corresponden a coberturas.

A lo largo de los últimos años ha aumentado el interés y la necesidad de disponer de una información de usos y coberturas del territorio fiable y actualizada, siendo numerosos los proyectos de carácter local, nacional e internacional cuyo objetivo es la creación y actualización de bases de datos de usos y ocupación del suelo.

Los procedimientos metodológicos para llevar a cabo estas tareas requieren un alto grado de intervención humana, ya que están basados casi íntegramente en la fotointerpretación de imágenes, o su comparación con otras imágenes u otra base de datos. Sin embargo las recientes mejoras en la calidad de los sensores, tanto aerotransportados como espaciales, han supuesto un incremento significativo en la disponibilidad de imágenes de alta resolución.

Entre las nuevas misiones satelitales destaca Sentinel-2 (S2). El lanzamiento del primer satélite S2 en 2015 por la ESA (*European Space Agency*) ha despertado un gran interés en la comunidad científica debido a sus múltiples aplicaciones, pues su alta resolución espacial, temporal, espectral y radiométrica, así como la distribución gratuita de los datos adquiridos abren una nueva era en la observación de la Tierra (ESA, 2015), mejorando las misiones operativas actuales, como pueden ser Landsat o SPOT.

El primer objetivo del presente trabajo es cuantificar la mejora en las clasificaciones con S2 respecto a otras misiones como SPOT, aplicando diferentes métodos con diferentes porcentajes de entrenamiento. El segundo objetivo es estudiar qué clasificador automático es mejor para clasificar las imágenes de S2, individualmente y combinando todos los clasificadores, para calcular la cantidad de cada cultivo que existe en una zona. Usaremos para ello datos de campo e imágenes de dos zonas agrarias diferentes: una en la huerta cercana a Valencia (España), y otra en una zona agraria del sur de Argentina.

2. Métodos

En primer lugar mostraremos las diferentes técnicas de clasificación de imágenes seleccionadas y posteriormente los indicadores de calidad usados en este trabajo.

2.1. Técnicas de clasificación

La clasificación supervisada parte de un grupo de elementos pertenecientes a la imagen, conocidos como áreas de entrenamiento, de los que se conoce con un nivel de exactitud alto la clase a la que pertenecen. La clasificación del conjunto de la imagen es el proceso por el que a cada elemento contenido en la imagen se le asigna una categoría, basándose en los atributos contenidos en las áreas de entrenamiento. La clasificación supervisada fuerza el resultado para que se corresponda con coberturas del suelo definidas por el usuario, y por tanto, de su interés, pero no garantiza que las clases sean separables desde el punto de vista estadístico (Rees, 2005).

Debido al interés de la clasificación, se han desarrollado un elevado número de clasificadores automáticos que pueden ser divididos en cuatro grupos (Hastie *et al.*, 2009). En este trabajo se ha elegido un clasificador de cada grupo, seleccionando los más utilizados: un clasificador paramétrico lineal, un clasificador no paramétrico que organiza de forma jerárquica el conjunto de decisiones, un clasificador no paramétrico que depende de los vecinos más cercanos y un clasificador no paramétrico basado en árboles de decisión. En concreto los métodos seleccionados son LDA (*Linear Discriminant Analysis*) (Camps-Valls

et al., 2008), TREE (Árbol de decisión) (Breiman, 1984), K-NN (Camps-Valls, 2016), y RF (*Random Forest*) (Breiman, 2001).

LDA o *Linear Discriminant Analysis* es una técnica de clasificación basada en el aprendizaje supervisado. La idea central de LDA es obtener una proyección de los datos entrantes, con el fin de que la separabilidad de las clases sea la mayor posible. Cuando las clases están bien separadas, las estimaciones de parámetros de modelos de regresión logística son sorprendentemente inestables, en cambio el LDA no sufre de este problema (James *et al.*, 2015). Es una técnica supervisada ya que para poder buscar esa proyección se debe entrenar el sistema con patrones etiquetados (Camps-Valls *et al.*, 2008). El modelo de análisis discriminante utiliza un algoritmo que busca funciones o vectores discriminantes, esto es, combinaciones lineales o cuadráticas de las variables originales que maximizan la varianza de las distribuciones entre clases a la vez que minimizan las varianzas inter-clases. Una vez determinada la función discriminante es posible establecer fronteras explícitas que separan las regiones del espacio que pertenecen a cada categoría, mediante una recta, o un plano o hiperplano, donde se igualan las probabilidades de pertenencia a una u otra clase. Esta línea recta o plano es ortogonal a la función discriminante encontrada. LDA es un clasificador con un buen historial de proyectos debido a que los datos sólo pueden soportar límites de decisión simples, tales como lineal o cuadrático, y las estimaciones proporcionadas a través de los modelos gaussianos son estables (Hastie *et al.*, 2009).

Los árboles de decisión (TREE) son métodos no paramétricos en clasificación. Un árbol representa un conjunto de condiciones que se organizan de forma jerárquica, y que se aplican sucesivamente desde una raíz hasta llegar a un nodo terminal u hoja del árbol (Breiman, 1984; Quinlan, 1993). Los árboles de decisión se representan mediante un gráfico con estructura arbórea que ha sido inducido desde un conjunto de datos de entrenamiento. Todo árbol de decisión comienza con un nodo al que pertenecen todos los casos de la muestra que se quiere clasificar. A este nodo se le denomina raíz. Además del nodo o raíz, en un árbol de decisión pueden distinguirse nodos internos y nodos terminales, que también se denominan hojas. Los nodos internos se corresponden con

una condición elegida en el conjunto de atributos y cada ramificación de un nodo interno representa un valor del dominio de la condición del atributo. Los nodos terminales u hojas contienen una única etiqueta, que corresponde, en este caso, a la cubierta del suelo. Desde la raíz del árbol hasta una determinada hoja, el camino conforma una regla de clasificación (Rodríguez-Galiano *et al.*, 2012). Si el objetivo final es un resultado de clasificación tomando valores desde 1 hasta las k clases, los únicos cambios necesarios en el algoritmo del árbol pertenecen a los criterios para seleccionar nodos (Hastie *et al.*, 2009). Los árboles de decisión predicen que cada observación pertenece a la clase de entrenamiento más frecuente en dicha región de observaciones (James *et al.*, 2015).

El algoritmo de los k vecinos más próximos (K-NN) es un método no paramétrico en clasificación y regresión. Estima la clase de un individuo (píxel, parcela, etc.) en función de los “ k ” vecinos más cercanos mediante un sistema de selección, donde la mayoría determina la clase, es decir, cada elemento de entrenamiento tiene una etiqueta de clase asociada, y la clasificación de un punto dependerá del elemento de entrenamiento que tenga más cercano. El más cercano es generalmente definido por la distancia Euclídea en el espacio de características, después de que cada característica se haya normalizado para tener la media 0 y la varianza 1 en el dato de entrenamiento, ya que es posible que tengan diferentes unidades. Además la distancia Euclídea es apropiada para las características cuantitativas. El modelo K-NN está basado en la memoria, y no requiere ningún otro modelo para estar entrenado. A pesar de su simplicidad, K-NN suele ser exitoso cuando cada clase tiene muchos elementos de entrenamiento y el límite de decisión es muy irregular (Hastie *et al.*, 2009).

El algoritmo Random Forest (RF) que fue propuesto por Breiman (2001), construye una gran colección de árboles des-correlacionados, y luego los promedia. Según James *et al.* (2015), los Random Forest proporcionan una mejora respecto a los árboles de decisión por el hecho de construir estos árboles des-correlacionados. Se trata de un algoritmo no paramétrico de aprendizaje automático ampliamente utilizado, que genera y combina un conjunto de árboles de decisión, donde cada uno se genera usando un subconjunto de elementos de entrenamiento diferente en cada iteración,

y los nodos se dividen utilizando la mejor de las variables de una muestra aleatoria de estas. La clase de salida se obtiene como la selección de la mayoría de las salidas de un gran número de árboles individuales (Breiman, 2001), es decir, cada árbol ha votado la clase que consideraba, y luego se ha clasificado usando el voto de la mayoría (Hastie *et al.*, 2009). El algoritmo produce una estimación objetiva interna de la generalización de error, utilizando las llamadas “out-of-bag”, el cual es un método para medir el error de predicción de los Random Forest cuando existen muestras que no están incluidas en el subconjunto de entrenamiento. El muestreo aleatorio conduce a una mayor estabilidad y una mejor exactitud de la clasificación en comparación con un enfoque único de un árbol de decisión. Este clasificador también es relativamente insensible a la cantidad de datos de entrada y multicolinealidad de los datos (Hastie *et al.*, 2009). RF se ha aplicado con éxito en varios problemas de regresión y clasificación de los datos de observación de la Tierra y en general logra buenos resultados (Gislason *et al.*, 2006).

Inicialmente se puede suponer que los cuatro algoritmos seleccionados serán bastante precisos y de su aplicación se obtendrán resultados aceptables. Pero además, en este trabajo se pretende demostrar que ningún algoritmo de clasificación independiente es el óptimo y que es posible obtener mejores resultados por la combinación de dos o más algoritmos. La combinación de clasificadores se realizó mediante máximo voto (es decir, cada píxel es clasificado con todos los métodos y al final se asigna a la clase que más veces se repite) por resultar muy sencilla de implementar y robusta en general.

2.2. Indicadores de calidad de los clasificadores

Un tema clave es la evaluación de fiabilidad de los procedimientos empleados y su calidad en los resultados finales. La calidad final de un método de clasificación o de combinación de clasificadores, se puede extraer a partir de un coeficiente de precisión, de forma que se otorga mayor confianza a aquel clasificador que haya demostrado mayor acierto. En este trabajo usaremos la matriz de confusión y derivaremos distintos índices de calidad a partir de ella.

La matriz de confusión, también llamada *tabla de contingencias*, consta de una tabla de doble entrada, que confronta los valores reales o *verdad-terreno* con los resultados de la clasificación. La diagonal de la matriz nos muestra la cantidad de píxeles reales y de la clasificación que coinciden por categoría, mientras que los restantes, nos vienen a decir aquellos que se confunden con otras categorías. Es decir, en sentido vertical se representa el porcentaje de píxeles reales que se confundieron en la verdad-terreno, y en sentido horizontal el porcentaje de píxeles después de la clasificación que se han confundido.

Como medida de precisión, se ha utilizado el índice Kappa, propuesto por Cohen (1960). El índice Kappa se usa para evaluar la concordancia de métodos cuyo resultado es categórico, con dos o más clases. Este índice representa la proporción de acuerdos observados respecto del máximo acuerdo posible más allá del azar. En la interpretación del índice Kappa hay que tener en cuenta que el índice depende del acuerdo observado, pero también de la prevalencia del carácter estudiado y de la simetría de los totales marginales (Abraira, 2001).

Para cuantificar la mejora de S2 respecto a SPOT, calcularemos el índice Kappa con todas las bandas de S2 y con las de SPOT, usando diferentes porcentajes de entrenamiento (10, 20, 30, 40 y 50% de datos para entrenamiento y el resto de validación). La división entre datos de entrenamiento y validación se realiza al azar. Para cada porcentaje se han aplicado los cuatro clasificadores, y la combinación de ellos. Además calcularemos el intervalo de confianza para un nivel de significación del 95%, para determinar cómo la mejora obtenida es estadísticamente significativa.

3. Datos

Los datos de los que se parte en este trabajo son por una parte los datos de campo, obtenidos de la información sobre el terreno de las dos zonas en estudio, y por otra parte las imágenes de teledetección de S2.

3.1. Datos de campo

Las dos zonas en estudio son dos áreas distantes entre sí, una situada en el continente europeo y la

otra en el americano. La primera corresponde a la zona de huerta ubicada entre Valencia, Burjassot, Moncada y Alboraya, situada en la zona norte del área metropolitana de Valencia (España), en la comarca de l'Horta Nord (Figura 1).

La superficie es casi llana y se encuentra a unos 50 m de altitud sobre el nivel del mar. Como toda la comarca, posee un clima mediterráneo. En ella se cultivan principalmente cítricos y hortalizas. En cada parcela se pueden obtener 2 o 3 cosechas al año de diferentes cultivos, por lo que en cada época del año se observan diferentes especies vegetales. Con una extensa red de acequias construidas en el periodo islámico, se canaliza el agua del río Turia para llevar a cabo el regadío que se realiza por inundación. Las parcelas suelen tener entre 50 y 100 m aproximadamente de lado.

La segunda área de estudio se corresponde al Valle Bonaerense del Río Colorado (VBRC), en Argentina. La región se ubica al sudoeste de la provincia de Buenos Aires sobre los partidos de Villarino y Patagones (Figura 2). Está formada por las localidades de Pedro Luro, Hilario Ascasubi, Villalonga, Juan A. Pradere y Mayor Buratovich. Se encuentra ubicada aproximadamente a 100 km al sur de la ciudad de Bahía Blanca y a 800 km de la Capital Federal, Buenos Aires.

El VBRC tiene un clima semiárido estepario templado. Los suelos son muy variados con predominio de los suelos arenosos que permiten el desarrollo de una amplia gama de cultivos,

requiriendo un buen sistema de drenaje para evitar los problemas de salinidad originados por la desembocadura del Río Curacó sobre el Río Colorado. El aporte de sales de esta agua al Colorado es una enorme preocupación para el Valle Bonaerense del Río Colorado. El principal cultivo de esta zona es la cebolla, complementado con girasol, maíz y alfalfa, entre otros cultivos (Pezzola, 2014).

En la zona de estudio de la huerta de Valencia, se tomaron los datos los días 1 y 2 de Mayo de 2016, coincidiendo con una imagen de S2 despejada de la zona (del 1 de Mayo). Para el trabajo se seleccionaron y etiquetaron 14 clases con un total de 91 parcelas con diferentes cultivos o suelos: 2 parcelas de suelo abandonado, 11 parcelas de suelo desnudo, 4 parcelas de alfalfa (*Medicago sativa*), 9 parcelas de asfalto, 11 parcelas de cauce seco, 6 parcelas de cebolla (*Allium cepa*), 1 parcela de coles (*Brassica oleracea*), 3 parcelas de lechuga (*Lactuca sativa L.*), 9 parcelas de naranjo (*Citrus sinensis*), 1 parcela de palmeras (*Phoenix roebelenii*), 13 parcelas de



Figura 1. Mapa de ubicación de la huerta de Burjassot, València.



Figura 2. Mapa de ubicación del VBRC, Buenos Aires, Argentina.

patata (*Solanum tuberosum*), 9 parcelas de construcciones, 3 parcelas de viveros y 9 parcelas de alcachofa (*Cynara scolymus*). En cada una de las 91 parcelas se ha tomado un ROI (*region of interest*) de entre 12 y 16 píxeles centrales, de manera que se dispone al final de 1387 píxeles como base de datos de entrenamiento y validación.

En el estudio del Valle Bonaerense del Río Colorado, la toma de los datos se realizó en el mes de Marzo los días 2 y 17. En la campaña argentina fueron seleccionadas y etiquetadas 11 clases con un total de 190 parcelas con diferentes cultivos o suelos: 41 parcelas de agua, 9 parcelas de viveros, 13 parcelas de caminos, 14 parcelas de construcciones, 17 parcelas de pasto natural (*hierba*), 16 parcelas de maíz (*Zea mays*), 3 parcelas de girasoles (*Helianthus annuus*), 23 parcelas de cebolla (*Allium cepa*), 17 parcelas sin cultivo, 23 parcelas de alfalfa (*Medicago sativa*), y 14 parcelas de agropiro (*Thinopyrum ponticum*). La imagen S2 usada para clasificar los cultivos fue seleccionada el 16 de Enero, pues en esa época los cultivos se encontraban en pleno desarrollo. En cada una de las 190 parcelas se ha tomado un ROI de entre 20 y 50 píxeles centrales, pues las parcelas son mucho más grandes que en la huerta de Valencia, de manera que se dispone al final de 8032 píxeles como base de datos de entrenamiento y validación.

En las dos áreas de estudio se han usado los datos de campo como datos de entrenamiento y validación, realizando las particiones de entrenamiento y validación-test mediante asignación aleatoria de los píxeles. El hecho de que el número de píxeles de cada base de datos sea superior a 1200 asegura que al dividirlos aleatoriamente en dos grupos tengamos píxeles de las distintas clases distribuidos por toda la zona. Sin embargo, solo en el área de la huerta de Burjassot-Valencia se hizo el estudio comparando los cuatro métodos y cuantificando la mejora respecto a SPOT. Una vez que se ha determinado que el mejor clasificador es el K-NN se ha aplicado este método a ambas imágenes en las dos zonas de estudio.

3.2. Descripción de datos SENTINEL-2

Las imágenes usadas en este trabajo fueron tomadas por S2, misión que forma parte del Programa Copérnico para la Observación de la Tierra de la

ESA. Este programa fue diseñado para suministrar información actualizada y de fácil acceso para mejorar la gestión del medio ambiente, y comprender y mitigar los efectos del cambio climático. El programa tiene planificadas cinco misiones y cada una está conformada por varios satélites.

La misión S2 se basa en una constelación de dos satélites. El primero, S2A, fue lanzado el 23 de junio de 2015. El S2B fue lanzado el 7 de marzo de 2017. El par de satélites S2 entregará datos de todas las superficies de la Tierra con un tiempo efectivo de revisita en el Ecuador de cinco días.

S2 proporciona datos para la generación de productos tales como: mapas de usos del suelo, mapas de cambios a lo largo del tiempo, mapas de riesgo, e imágenes rápidas para prevención de desastres. También proporciona datos para la generación de variables biofísicas de la vegetación como la cobertura vegetal, el contenido de clorofila o el contenido de agua en las hojas. Este satélite cuenta con un instrumento de captura multispectral mejorando los datos de otras misiones operativas (Drusch et al., 2012).

El instrumento multispectral (MSI) a bordo de S2 cuenta con 13 bandas espectrales, las cuales van desde el espectro visible y el infrarrojo cercano (NIR), hasta el infrarrojo de onda corta (SWIR). Entre ellas hay cuatro bandas de 10 m (tres del visible y una del NIR), seis bandas de 20 m, y tres bandas de 60 m de resolución para corrección atmosférica y detección de nubes.

S2 incorpora tres nuevas bandas en la región del *red-edge* que mejora la configuración del sensor para estudios de aguas y vegetación (Delegido et al., 2011). Aunque está previsto que S2 proporcione productos de alto nivel (Nivel 2b/3), actualmente los datos S2 disponibles son un producto de radiancia TOA formado por un conjunto de grillas (*granules*) contiguas de 100 km², correspondientes al nivel 1c. Para el procesamiento de los datos 1c la ESA ha desarrollado el software Sen2Cor (Sentinel 2 Corrections), basado en un conjunto de técnicas para la corrección atmosférica y un módulo de clasificación de escena (ESA, 2016). El resultado es un producto de reflectividades BOA, correspondiente al nivel 2a, además de un mapa de espesor óptico de aerosoles, un mapa de vapor de agua y un mapa de clasificación de la escena junto con los indicadores de calidad.

Sen2Cor se basa en una amplia base de *look-up tables* que son compiladas a partir del modelo de transferencia radiativa atmosférica basada en libRadtran1.

El satélite S2 mejora misiones operativas como Landsat o SPOT proporcionando bandas más estrechas y con canales adicionales para evaluar vegetación, y bandas dedicadas a la mejora de la corrección atmosférica y la detección de nubes cirrus. La órbita de este satélite es heliosincrónica y se encuentra a 786 km de altitud, tomando la imagen de una zona a las 10:30 am. Esta hora local es la mejor para reducir al mínimo la cobertura de nubes y asegurar una iluminación del Sol plena, y además está cerca de la hora de paso de Landsat y casi idéntica a la del SPOT-5. El satélite está diseñado para una vida útil de 7 años, con propulsor de 12 años de operaciones, incluyendo la salida final de órbita.

La misión está dedicada a una cobertura plena de las superficies terrestres (incluyendo las islas mayores) a partir de los 56° S (América del Sur) hasta los 83° N (Groenlandia).

Para la obtención de datos, la ESA ha desarrollado un portal en el que se tiene acceso gratuitamente a las imágenes del satélite Sentinel-2, pudiendo acceder también a las imágenes de Sentinel-1. Las imágenes utilizadas son del día 1 de Mayo de 2016 para Valencia, y del 13 de Enero de 2016 para Argentina. Fueron elegidas por no tener nubes y fueron corregidas atmosféricamente con Sen2cor.

4. Resultados

En esta sección se mostrarán los resultados de los índices Kappa calculados con todos los clasificadores para ambas imágenes. Después obtendremos la matriz de confusión para analizar los resultados obtenidos. Posteriormente se mostrarán los mapas de usos del suelo obtenidos con el mejor clasificador para ambas zonas. Y por último, se calculará el área ocupada por cada cultivo en las zonas de estudio.

4.1. Índice Kappa

En la Tabla 1 se muestran los índices Kappa para la imagen de España, aplicando todos los clasificadores mencionados, y los porcentajes de entrenamiento entre el 10 y el 50%, para una

imagen con todas las bandas S2, y en la Tabla 2 con solo las cuatro bandas de S2 que coinciden con SPOT (que son las 3 del visible y la del NIR). Tal como se aprecia en la Tabla 1, cuando usamos todas las bandas, todos los clasificadores han dado un resultado bueno, pero algunos han dado incluso mejor que la combinación, como son el K-NN y el RF.

Tabla 1. Índice Kappa para la imagen de Burjassot con todas las bandas S2, con porcentajes de píxeles aplicados a entrenamiento quedando el resto como validación.

%	LDA	K-NN	Tree	RF	Comb.	Comb.	Comb.
						Sin LDA	ni Tree
10	0,76	0,82	0,62	0,84	0,86	0,82	0,84
20	0,81	0,88	0,75	0,88	0,89	0,88	0,89
30	0,80	0,91	0,81	0,91	0,91	0,91	0,92
40	0,81	0,94	0,82	0,94	0,92	0,93	0,94
50	0,80	0,96	0,87	0,95	0,95	0,96	0,96

Si analizamos por partes las Tablas 1 y 2, observamos que el clasificador LDA, nos da un resultado Kappa que corresponde a un grado de acierto sustancial, tanto con 4 bandas como con todas.

Tabla 2. Índice Kappa para la imagen de Burjassot con 4bandas SPOT (RGB y NIR), con porcentajes de píxeles aplicados a entrenamiento quedando el resto como validación.

%	LDA	K-NN	Tree	RF	Comb.	Comb.	Comb.
						Sin LDA	ni Tree
10	0,70	0,79	0,61	0,77	0,78	0,77	0,79
20	0,71	0,85	0,75	0,82	0,83	0,83	0,85
30	0,71	0,89	0,78	0,88	0,86	0,88	0,90
40	0,72	0,91	0,79	0,90	0,89	0,90	0,91
50	0,72	0,94	0,82	0,92	0,91	0,92	0,94

El clasificador K-NN da unos resultados con un grado de acierto casi perfecto en todas las clasificaciones, excepto con el 10% de entrenamiento con la imagen de cuatro bandas, que da un índice de 0,79, considerado sustancial con el acierto. Además empieza a dar mejores resultados que la combinación a partir del 40% de datos de entrenamiento con todas las bandas, mientras que con cuatro bandas, da siempre mejor que la combinación.

El clasificador Tree, cuando se clasifica con todas las bandas, da un grado de acierto sustancial para

el 10 y el 20% de entrenamiento, y a partir del 30% tiene un grado de acierto casi perfecto. Con cuatro bandas el resultado es más desfavorable, con lo que el grado de acuerdo es sustancial hasta el 40% de entrenamiento, y solo con el 50% se puede considerar casi perfecto. Este clasificador siempre proporciona peores resultados que la combinación.

El clasificador RF da un resultado de acierto casi perfecto con las bandas S2, y también con las cuatro bandas SPOT, excepto con el 10% de entrenamiento. A partir de 30% da mejor resultado que la combinación.

Tras este primer análisis ya se puede observar cómo los resultados obtenidos con todas las bandas son mejores que con cuatro bandas. Pero debido a que se quería demostrar que la combinación debía ser mejor que cualquiera de ellos, y se comprobó que no siempre se cumplía, se repitió la clasificación eliminando el peor de ellos, el cuál es el LDA, tanto con todas las bandas como con cuatro.

En la séptima columna de las Tablas 1 y 2 mostramos nuevamente todos los clasificadores, pero con la combinación realizada sin el clasificador LDA. En la clasificación con todas las bandas, vemos cómo la combinación ha mejorado respecto a todos los demás clasificadores, pero todavía hay porcentajes en los que queda por debajo de algunos de los clasificadores. Por ejemplo el clasificador K-NN a partir del 40% arroja mejores resultados que la combinación, y el RF también, mientras que en la clasificación con cuatro bandas se aprecia cómo la combinación ya da mejores resultados que el clasificador RF, pero en cambio el clasificador K-NN sigue siendo mejor que la combinación en todo momento.

Como todavía existía resultados en los que la combinación de clasificadores daba resultados más bajos que los clasificadores individualmente,

se repitió la clasificación con la combinación, pero esta vez eliminando de la combinación también el clasificador Tree, ya que después del LDA, era el que peores resultados proporcionaba. En la octava columna de las Tablas 1 y 2 tenemos los resultados de la combinación K-NN y RF, es decir, sin el clasificador LDA y el Tree.

Tras haber eliminado los dos peores clasificadores, se observa que la combinación ha mejorado tanto con todas las bandas como con cuatro, respecto a los clasificadores K-NN y RF. En algunos porcentajes la combinación supera a los clasificadores, y en otros los iguala, aunque si añadiésemos decimales se podría comprobar cómo ha mejorado ligeramente.

En la Tabla 3 se muestra la mejora obtenida en la clasificación realizándola con todas las bandas de S2 respecto a realizarla con solo las cuatro que coinciden con SPOT. Se puede observar que S2 permite mejorar la clasificación realizada con SPOT entre el 2 y el 5%, usando la combinación K-NN y RF. En las Tablas 4 y 5 se muestra cómo el intervalo de confianza da unos valores estadísticamente significativos.

Tabla 3. Cuantificación de la mejora de S2 respecto SPOT, con porcentajes de píxeles aplicados a entrenamiento quedando el resto como validación.

%	LDA	K-NN	Tree	RF	Comb. sin
					LDA ni Tree
10	0,06	0,03	0,01	0,06	0,05
20	0,10	0,03	0,01	0,06	0,04
30	0,09	0,02	0,03	0,03	0,02
40	0,09	0,02	0,03	0,03	0,03
50	0,08	0,02	0,05	0,03	0,02

4.2. Mapas de usos del suelo

En la Figura 3 se muestra el mapa de usos del suelo para la huerta de Burjassot en Valencia (España). El mapa se ha construido mediante el clasificador

Tabla 4. Intervalos de confianza para el 95% sobre la imagen de Burjassot con todas las bandas de S2, con porcentajes de píxeles aplicados a entrenamiento quedando el resto como validación.

%	LDA		K-NN		Tree		RF		Comb. sin LDA ni Tree	
10	0,76	0,83	0,82	0,86	0,63	0,68	0,84	0,87	0,84	0,87
20	0,81	0,85	0,84	0,88	0,76	0,80	0,87	0,91	0,89	0,92
30	0,80	0,84	0,91	0,94	0,81	0,85	0,90	0,93	0,91	0,94
40	0,81	0,85	0,93	0,95	0,82	0,86	0,93	0,95	0,93	0,96
50	0,80	0,84	0,95	0,97	0,87	0,90	0,95	0,97	0,95	0,97

Tabla 5. Intervalos de confianza para el 95% sobre la imagen de Burjassot con las 4 bandas de SPOT (RGB y NIR) , con porcentajes de píxeles aplicados a entrenamiento quedando el resto como validación.

%	LDA		K-NN		Tree		RF		Comb. sin LDA ni Tree	
10	0,70	0,75	0,79	0,84	0,62	0,67	0,78	0,82	0,72	0,77
20	0,71	0,76	0,85	0,89	0,75	0,80	0,82	0,86	0,83	0,87
30	0,72	0,76	0,89	0,92	0,79	0,83	0,88	0,91	0,88	0,91
40	0,73	0,77	0,91	0,94	0,79	0,83	0,90	0,93	0,90	0,93
50	0,73	0,77	0,93	0,96	0,82	0,86	0,92	0,94	0,92	0,95

K-NN, ya que era el mejor clasificador, y con un 50% de entrenamiento, dando un índice Kappa de 0,94.

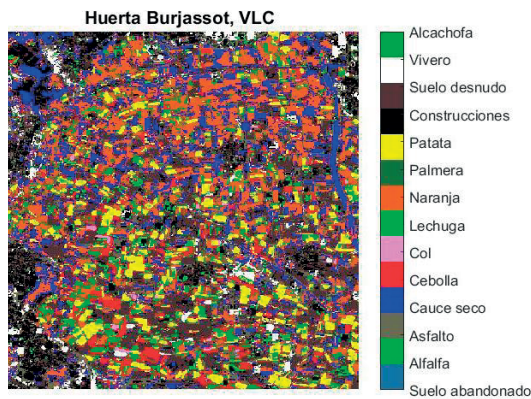


Figura 3. Mapa de usos del suelo de la huerta de Burjassot, Valencia.

En la Figura 4 se muestra el mapa de usos del suelo del Valle Bonaerense del Río Colorado, usando también el clasificador K-NN. La clasificación se ha realizado con el 50% de datos como entrenamiento, y ha dado como grado de acierto un índice Kappa de 0,83, lo cual se considera casi perfecto.

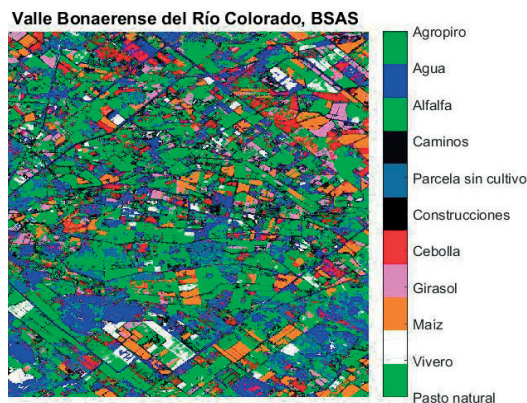


Figura 4. Mapa de usos del suelo del VBRC, BsAs.

4.3. Matriz de confusión

En la Tabla 6 se muestra la matriz de confusión para la zona de estudio de España, con las catorce categorías representadas, con la cantidad de los píxeles para cada clase en porcentaje.

Tras calcular la precisión global se obtuvo que el 96,52% de las muestras estaban correctamente clasificadas. Si analizamos la tabla 6 se aprecia cómo las clases mejor etiquetadas y clasificadas han sido el *suelo abandonado*, la *alfalfa* y los *viveros*, obteniéndose un 100% de aciertos. Por el contrario, la clase peor etiquetada y por lo tanto peor clasificada ha sido la *palmera* con un 63,63% de aciertos, por el hecho de tener pocos píxeles, y pocos datos de campo. El resto de clases presenta unos valores coherentes dentro de una clasificación.

En la Tabla 7 se muestra la matriz de confusión para la zona de estudio en Argentina, con las once categorías elegidas y con el porcentaje de la cantidad de píxeles para cada clase. Tras calcular la precisión global se obtuvo que el 85,28% de las muestras estaban correctamente clasificadas. Este valor fue inferior al valor de la imagen de España (96,52%), lo que nos llevó a pensar que existían varios factores que podrían haber afectado a este resultado: el área a clasificar de Argentina era mucho más grande que la de España, con lo que se le estaban inyectando más datos a los clasificadores; además había más tipos de clases en España que en Argentina, por lo que se podían diferenciar más cultivos; la toma de datos coincidía con el paso de Sentinel, mientras en Argentina no, porque, aunque los datos de campo fueron tomados en el mes de Marzo, se decidió trabajar con la imagen de Enero debido a que los cultivos estaban en pleno desarrollo. Otro factor podría ser el hecho de que el mejor método de clasificación en la huerta de Valencia, no lo sea para la zona de Argentina.

Tabla 6. Matriz de confusión de la clasificación para la huerta de Burjassot.

		Valores reales (en porcentaje)													
		Suelo aband.	Alfalfa	Asfalto	Cauce seco	Cebolla	Coles	Lechuga	naranja	Palmera	Patata	Construcciones	Suelo desn.	Vivero	Alcachofa
Resultados clasificación	CLASE														
	Suelo aband.	100,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Alfalfa	0,0	100,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Asfalto	0,0	0,0	96,7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,7	0,0	0,0	0,0
	Cauce seco	0,0	0,0	0,0	94,2	0,0	0,0	0,0	3,5	0,0	0,0	0,0	0,5	0,0	0,0
	Cebolla	0,0	0,0	0,0	1,9	97,9	0,0	0,0	0,0	9,1	0,4	0,0	0,5	0,0	3,6
	Coles	0,0	0,0	0,0	0,0	0,0	92,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Lechuga	0,0	0,0	0,0	0,0	0,0	0,0	94,7	0,0	0,0	0,0	0,0	0,5	0,0	0,0
	Naranja	0,0	0,0	0,0	1,3	0,0	0,0	0,0	93,0	9,1	0,0	0,0	0,0	0,0	0,0
	Palmera	0,0	0,0	0,0	1,3	0,0	0,0	0,0	1,2	63,6	0,0	0,0	0,0	0,0	0,0
	Patata	0,0	0,0	0,0	0,0	0,0	0,0	5,3	0,0	0,0	99,6	0,0	0,0	0,0	0,9
	Construc.	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	97,1	0,5	0,0	0,0
	Suelo desn.	0,0	0,0	3,3	1,3	0,0	0,0	0,0	0,0	0,0	0,0	2,1	98,1	0,0	0,0
	Vivero	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	100,0	0,0
	Alcachofa	0,0	0,0	0,0	0,0	2,1	8,0	0,0	2,3	18,2	0,0	0,0	0,0	0,0	95,5

Si analizamos la Tabla 7 se aprecia como la clase mejor etiquetada y clasificada ha sido el *maíz*, con un 95,1%. Por el contrario, la clase peor etiquetada y por lo tanto peor clasificada ha sido el *agropiro* con un 61,9% de aciertos.

4.4. Área ocupada por clase

En la Tabla 8 se muestran las cantidades de cultivo calculadas para cada zona asignando a cada píxel un área de 100 m². En la tabla de la huerta de Burjassot se aprecia cómo el cultivo que mayor

Tabla 7. Matriz de confusión de la clasificación para el VBRC, Argentina.

		Valores real(en porcentajes)										
CLASE		Pasto natural	Viveros	Maiz	Girasol	Cebolla	Construcciones	Parcela sin cultivo	Caminos	Alfalfa	Agua	Agropiro
Resultados de la clasificación	Pasto natural	84,8	0,0	0,0	0,1	2,0	6,4	6,1	0,5	21,0	0,3	0,0
	Viveros	0,0	93,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
	Maíz	0,0	0,0	95,1	1,8	0,0	0,9	0,0	0,0	0,0	0,0	11,9
	Girasol	0,0	0,0	0,0	87,7	0,0	0,0	0,0	1,5	0,0	0,0	0,0
	Cebolla	0,0	0,0	1,3	0,0	77,2	2,8	0,0	0,0	0,0	7,1	15,5
	Construcciones	0,0	0,0	0,0	0,1	0,0	70,6	0,0	1,5	0,0	1,8	1,8
	Parcela sin cultivo	1,0	0,0	0,0	0,0	8,3	1,8	90,9	1,0	4,8	1,6	0,0
	Caminos	0,3	6,9	0,0	0,0	0,0	6,4	2,9	76,4	76,4	0,0	0,0
	Alfalfa	13,9	0,0	0,1	6,0	0,0	1,8	0,0	0,5	74,1	74,1	3,4
	Agua	0,0	0,0	2,8	0,0	11,7	1,8	0,0	10,8	10,8	88,7	5,4
	Agropiro	0,0	0,0	0,7	4,2	0,7	7,3	0,0	7,9	0,0	0,2	61,9

Tabla 8. Área ocupada por las clases de la huerta de Burjassot, VLC, y VBRC, BsAs.

Huerta de Burjassot		INTA Argentina	
Clase	Área (ha)	Clase	Área (ha)
Suelo aband.	8,26	Parcela sin cult.	2665,74
Alfalfa	7,13	Girasol	804,47
Asfalto	185,79	Maíz	1210,27
Cauce seco	227,69	Cebolla	1073,30
Cebolla	101,80	Agua	1276,88
Col	10,86	Pasto natural	748,76
Lechuga	32,29	Agropiro	3309,03
Naranja	261,71	Alfalfa	2120,34
Palmera	14,46	Construcciones	1239,00
Patata	122,01	Caminos	1887,38
Construcc.	144,13	Viveros	2099,71
Suelo desn.	324,23		
Vivero	46,78		
Alcachofa	124,56		

superficie ocupa es el *naranja*, y el que menor ocupa es la *alfalfa*. En la tabla de la zona de estudio de Argentina se observa cómo el cultivo que más área ocupa es el *agropiro*, mientras que el *pasto natural* es el que menos ocupa.

5. Conclusiones

En este trabajo se han obtenido los mapas de usos del suelo con el área de cada cultivo para dos zonas agrarias con imágenes S2. Se ha observado que el mejor clasificador para imágenes S2 es el K-NN, y la combinación del K-NN y el RF. Además, se ha demostrado la mejora de S2 respecto a SPOT cuantificándose entre 2 y 5% por encima. Por otro lado, la precisión global para la huerta de Burjassot fue 96,52%, mientras que para la zona de Argentina fue del 85,28%, es decir, el porcentaje de que las muestras estuviesen correctamente clasificadas fue mejor en España que en Argentina. Finalmente, los resultados de este trabajo muestran que la mejora en la resolución espacial y radiométrica de S2 respecto a otras misiones operativas, permite también la mejora en los resultados de la clasificación.

Referencias

Abaira, V. 2001. El índice kappa. *Semergen*, 27(5), 247-249. [https://doi.org/10.1016/S1138-3593\(01\)73955-X](https://doi.org/10.1016/S1138-3593(01)73955-X)

- Breiman, L. 1984. Classification and regression trees. *Chapman & Hall/CRC*.
- Breiman, L. 2001. Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Camps-Valls, G., Gómez Chova, L., Muñoz Marí, J., Rojo Álvarez, J.L., Martínez-Ramón, M. 2008. Kernel based framework for multi-temporal and multi-source remote sensing data classification and change detection. *IEEE Trans. Geosc. Rem. Sens.*, 46(6), 1822-1835. <https://doi.org/10.1109/TGRS.2008.916201>
- Camps-Valls, G., Tuia, D., Gómez-Chova, L., Jiménez, S., Malo, J. 2011. Remote Sensing Image Processing. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 5(1), 1-192. <https://doi.org/10.2200/S00392ED1V01Y201107IVM012>
- Cohen J. A. 1960. Coefficient of agreement for nominal scales. *Educ. Psychol. Meas.*, 20(1), 37-46. <https://doi.org/10.1177/001316446002000104>
- Comber, A., Fisher, P., Wadsworth, R. 2005. You know what land cover is but does anyone else? ... an investigation into semantic and ontological confusion. *International Journal of Remote Sensing*, 26(1), 223-228. <https://doi.org/10.1080/0143116042000274032>
- Del Bosque, I., Arozarena, A., Villa, G., Valcárcel, N., Porcuna, A. 2005. Creación de un sistema de información geográfico de ocupación del suelo en España. "Proyecto SIOSE". *Actas del XI Congreso Nacional de Teledetección*, 21-23 Septiembre, Puerto de la Cruz, España, 255-262.
- Delegido, J., Verrelst, J., Alonso, L., Moreno, J. 2011. Evaluation of Sentinel-2 Red-Edge Bands for Empirical Estimation of Green LAI and Chlorophyll Content. *Sensors*, 11, 7063-7081. <https://doi.org/10.3390/s110707063>
- Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernández, V., Gascón, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., Spoto, F., Sy, O., Marchese, F., Bargellini, P. 2012. Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote sensing of Environment*, 120, 25-36. <https://doi.org/10.1016/j.rse.2011.11.026>
- ESA. 2016. www.esa.int/. Consultado el 10 de julio de 2016.
- Gislason, P.O., Benediktsson, J.A., Sveinsson, J.R. 2006. Random Forests for land cover classification. *Pattern Recognition Letters*, 27(4), 294-300. <https://doi.org/10.1016/j.patrec.2005.08.011>
- Hastie, T., Tibshirani, R., Friedman, J. 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed.; Springer: New York, NY, USA. <https://doi.org/10.1007/978-0-387-84858-7>

- Immitzer, M., Atzberger, C., Koukal, T. 2012. Tree species classification with Random Forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, 4, 2661-2693. <https://doi.org/10.3390/rs4092661>
- James, G., Witten, D., Hastie, T., Tibshirani, R. 2015. *An Introduction to Statistical Learning with Applications in R*. In Springer Texts in Statistics. Volume 103. <https://doi.org/10.1007/978-1-4614-7138-7>
- Landis, J.R., Koch, G.G. 1977. The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174. <https://doi.org/10.2307/2529310>
- López González, F.J., Crecente Maseda, R., Álvarez López, C.J. 2002. Los usos del suelo analizados mediante S.I.G. XIV Congreso de Ingeniería Gráfica. 5-7 Junio, Santander, España.
- Manual de usuario de Sentinel 2-ESA. Julio 2015.
- Martimort, P., Berger, M., Carnicero, B., Del Bello, U., Fernández, V., Gascón, F., et al. 2007. Sentinel-2: Optical high-resolution mission for GMES operational services. *Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International*, 131, 18-23. <https://doi.org/10.1109/igarss.2007.4423394>
- Mena, A.J. 2014. Procesamiento de imágenes satelitales multiespectrales. *Proyecto final de carrera, Facultad de Informática, Universidad del País Vasco*.
- Pezzola, A. 2014. Transformaciones territoriales producto del riego en el Valle Bonaerense del Río Colorado. *Departamento de geografía y turismo, UNS, provincia de Buenos Aires*.
- Quinlan, J.R. 1993. Programs for Machine Learning. *1st ed. San Mateo, CA, Morgan*.
- Rees, G. 2005. The Remote Sensing Data Book. *Cambridge University Press*, 262 pp.
- Rodríguez-Galiano, V., Chica-Rivas, M. 2012. Clasificación de imágenes de satélite mediante software libre: Nuevas tendencias en algoritmos de Inteligencia artificial. *Departamento de Geodinámica, Universidad de Granada*.