

## **Metodología de clasificación automática de uso y cobertura de suelo**

### **Uso de Métodos de Aprendizaje Automático y teledetección para clasificación de uso y cobertura del suelo en un valle semiárido de la Patagonia**

Magda Alexandra Trujillo-Jiménez<sup>1</sup>, Ana Laura Liberoff<sup>2</sup>, Natalia Pessacg<sup>2</sup>, Cristian Pacheco<sup>2</sup>, Silvia Flaherty<sup>3</sup>

- <sup>1</sup> Laboratorio de Ciencias de las Imágenes, Departamento de Ingeniería Eléctrica y Computación, Universidad Nacional del Sur, Argentina  
<sup>2</sup> Instituto Patagónico para el Estudio de los Ecosistemas Continentales CCT CONICET CENPAT, Argentina  
<sup>3</sup> Universidad Nacional de la Patagonia San Juan Bosco, Argentina

**Resumen.** El valle agrícola-ganadero en la Cuenca Inferior del Río Chubut (VIRCh) tiene una extensión de 225 km<sup>2</sup> y está situado en la Patagonia semiárida oriental argentina. Si bien este es uno de los valles más importantes de la región no existe hasta ahora información espacial completa, detallada y actualizada de sus actividades y coberturas del suelo. En este artículo avanzamos en la sistematización de este proceso y comparamos 7 métodos tradicionales de aprendizaje automático supervisado aplicados a la clasificación del uso y la cobertura del suelo a partir de imágenes satelitales de Sentinel-2 MSI y de datos adquiridos sobre el terreno. Con estos métodos obtuvimos predicciones que superan entre el 70 y 80 % de precisión en la clasificación de cultivos frutales, horticultura, terrenos construidos, arbustales, pasturas y agua. Nearest Neighbors y Decision Tree fueron los clasificadores que mostraron los mejores desempeños. El código fuente y algoritmos desarrollados están disponibles en: [DOI: 10.5281/zenodo.5338597](https://doi.org/10.5281/zenodo.5338597) para su libre recreación, uso y reproducibilidad.

**Palabras clave.** Aprendizaje Automático Supervisado, Uso y Cobertura del Suelo, Imágenes Satelitales Multiespectrales, Valle Inferior del Río Chubut.

## **1 Introducción**

Los mapas de uso y cobertura del suelo (LULC por sus siglas en inglés) son representaciones del territorio que permiten evaluar los estados pasados y presentes y analizar tendencias de cambio en las matrices socio-ambientales de un determinado

paisaje. Los mapas de LULC también son insumos clave para la evaluación y modelización espacial de los servicios ecosistémicos (e.g., [1], [2], [3]). Como tal son insumos básicos para el ordenamiento territorial y las políticas públicas. Producir mapas de LULC precisos es una tarea difícil que requiere una cantidad importante de datos adecuados, tanto de campo como a nivel de paisaje (es decir, teledetección) y algoritmos de clasificación eficientes capaces de combinar esos datos provistos por los sensores remotos y los adquiridos en el campo para caracterizar con precisión las coberturas del suelo presentes en un área en particular [6], [7], [8].

La teledetección cumple un rol esencial en el mapeo de LULC. Las imágenes multispectrales adquiridas por los satélites Landsat -disponibles desde principios de la década de 1970- han sido muy utilizadas para la clasificación de LULC con diferentes grados de éxito [6]. Sin embargo, el lanzamiento de los satélites Sentinel-2A y 2B de la ESA en 2015 y 2017 respectivamente, proporcionan imágenes con una resolución espacial más fina (10 y 20 m), intervalos de revisita más cortos (5 días) y mayor resolución espectral. Estas mejoras han permitido a la comunidad científica el libre acceso a imágenes de alta calidad especialmente diseñadas para el estudio de la vegetación, lo que ha permitido aumentar la precisión global de la clasificación de LULC -incluyendo la clasificación de cultivos- [10], [11], [12], al tiempo que se ha incrementado el volumen de datos y el coste computacional.

Para generar mapas de LULC a partir de la información que proporciona la teledetección existen diferentes métodos, algunos de ellos basados en clasificadores supervisados de aprendizaje automático [8], los cuales se han utilizado ampliamente para la clasificación basada en píxeles [13], [14], [15]. Las numerosas ventajas de estos métodos incluyen la capacidad de manejar un elevado número de variables de entrada y datos espectrales no distribuidos normalmente, y de cartografiar las clases de LULC en paisajes complejos. Esto se traduce en mejores rendimientos globales de clasificación en comparación con las técnicas tradicionales, como la distancia mínima y la máxima verosimilitud, que son aún ampliamente utilizadas y están disponibles en la mayoría de los paquetes de procesamiento de imágenes [8], [16], [17], [18].

El valle agrícola de la Cuenca Inferior del Río Chubut (VIRCh) se encuentra en una zona semiárida y es uno de los valles agrícolas más importantes de la región patagónica. Las actividades agrícolas y los centros urbanos que dependen del Río Chubut para la provisión de agua son fuentes reconocidas de impacto sobre la calidad del agua [4], [5] y aunque el agua superficial es escasa, no existe un plan formal de gestión del uso del agua [37]. El único mapa de LULC disponible fue producido manualmente con una combinación de inspección visual de fotografías aéreas y toma de datos de campo por una institución gubernamental local en el año 2015 (Com.Pers. Pedro Tagliabue, Compañía de Riego del VIRCh, datos sin publicar)

En este contexto, el desarrollo de una metodología para producir periódicamente mapas precisos de LULC es un elemento clave para supervisar los cambios en el paisaje, en la matriz de desarrollo y aportar elementos para el manejo integrado del agua, así como para promover futuras investigaciones en esta región. El objetivo de este trabajo es comparar siete de los métodos de aprendizaje automático supervisado

más usados para clasificar con precisión las categorías de LULC, especialmente los cultivos, y generar una metodología que permita producir regularmente mapas de LULC actualizados del valle agrícola del VIRCh.

## **2 Materiales y métodos**

### **2.1 Área de estudio y recolección de datos de campo**

El valle agrícola en estudio se ubica en una región semiárida de la Patagonia Argentina, caracterizada por un clima templado-árido con escasas precipitaciones (200 mm de lluvia media anual), altas tasas de evapotranspiración (más de 6 mm en un día de verano [37]), temperatura media anual de 13 °C y fuertes vientos (21 km h-1 de intensidad media anual). El Valle Inferior o Cuenca Inferior del Río Chubut (VIRCh) concentra cerca del 95% de la población total de la cuenca (276.000 habitantes) y el 50% de la actividad agrícola de la provincia del Chubut con una superficie regada de 22500 ha (Figura 1). Los principales cultivos de este valle son las pasturas, la horticultura y, en menor medida, los cultivos frutales. Otras actividades y usos del suelo importantes dentro del valle que también dependen de las aguas superficiales del Río Chubut son los emprendimientos de engorde de ganado, las industrias y las zonas urbanas.

Para este estudio definimos 6 clases principales de LULC que representan los cultivos más importantes (Pasturas, Horticultura y Cultivos Frutales), la cobertura vegetal natural más abundante (Arbustales), los usos del suelo (Construcciones) y los recursos naturales (Agua). Con el fin de recolectar datos de referencia de campo para el entrenamiento y validación de los métodos, se realizaron varias salidas de campo a lo largo de 2018 (junio, julio, agosto, septiembre, noviembre y diciembre) y del verano de 2019 (enero, febrero y marzo). Se tomaron fotos georreferenciadas en el campo y se registró el tipo de uso/cobertura del suelo de interés para cada foto. Las fotos georreferenciadas se utilizaron para crear un conjunto de datos etiquetados utilizando el software QGIS [19] y el complemento OpenLayers. Para cada punto estudiado se generó manualmente un polígono conteniendo una cobertura homogénea (es decir, que encerraba la misma categoría de LULC para todos los píxeles de la imagen de satélite dentro del polígono). Algunas categorías de LULC que eran fáciles de identificar en las imágenes de satélite (por ejemplo, construcciones, agua) se digitalizaron directamente basándose en imágenes de alta resolución de Google Earth y/o de satélite. Las capas de tipo polígono luego fueron transformadas a raster con un píxel de 10 m de resolución espacial para crear correspondencia con los datos de las imágenes satelitales (ver debajo) y entrenar los clasificadores. Para clasificar los principales cultivos, coberturas y actividades dentro del valle agrícola (Tabla 1) utilizamos imágenes Sentinel-2 MSI.



Figura 1 Ubicación y detalle del valle agrícola bajo estudio. La línea azul representa al Río Chubut.

**Tabla 1.** Definición de las clases de LULC

Clase	Descripción
Pasturas	Cultivos de gramíneas y/o alfalfa y/o pasturas mixtas en buen estado, mayormente utilizadas como forraje.
Horticultura	Cultivos de legumbres y hortalizas
Cultivos Frutales	Árboles frutales cultivados
Arbustales	Áreas con vegetación natural, áreas con vegetación regenerada
Construcciones	Áreas urbanizadas, construcciones rurales y construcciones industriales
Agua	Cuerpos de agua naturales o artificiales, curso principal del Río Chubut

## 2.2 Adquisición y preprocesamiento de imágenes Sentinel-2

La misión Sentinel-2 del programa Copernicus de la ESA consta de 2 satélites (2A y 2B). El instrumento multiespectral a bordo de los satélites Sentinel-2 (MSI, por sus siglas en inglés), incluye tres bandas espectrales en el “borde rojo” (Vegetation Red Edge) además de las bandas del espectro visible, infrarrojo cercano e infrarrojo de onda corta. Este instrumento fue diseñado para el monitoreo de la vegetación y el mapeo de la cobertura terrestre, entre otras aplicaciones [42].

Se necesitaron tres teselas (T19GGM, T19GG y T20GLT) de Sentinel-2A para cubrir el área de estudio. Las imágenes fueron adquiridas el 25 de enero de 2019 y se descargaron -de forma gratuita- del catálogo Copernicus [45]. Desde diciembre de 2018, la plataforma Copernicus provee los datos Sentinel-2 MSI procesados a nivel de reflectancia en superficie (BOA, Nivel 2A [42]). Las imágenes están codificadas en JPEG2000 y con un valor de cuantificación de 10.000. Es necesario aplicar un factor

de 1/10.000 a los valores de BOA para recuperar los valores físicos de reflectancia en superficie [38],[42].

Para este experimento y con el fin de evaluar el rendimiento de los métodos de clasificación LULC de aprendizaje automático, se excluyeron las tres bandas espectrales que proporcionan información sobre la composición atmosférica (es decir, Banda 1-aerosol/costero, Banda 9-vapor de agua y Banda 10 - cirros). Las seis bandas espectrales con 20 m de resolución espacial (Bandas 5, 6 y 7- vegetation red edge; Banda 8A- narrow NIR y Bandas 11 y 12- SWIR 1 y 2) se re-muestraron a 10 m y se combinaron con las cuatro bandas espectrales de 10 m (Banda 2- Azul, Banda 3- Verde, Banda 4-Rojo y Banda 8-NIR) para obtener así un archivo multi-capas de 10 bandas espectrales a 10 m de resolución espacial. Se generó un mosaico con los tres archivos de 10 bandas y posteriormente un recorte del área de estudio para reducir el tamaño del archivo y facilitar su manejo. Todo el procesamiento de las imágenes se llevó a cabo utilizando software gratuito de SIG (QGIS v 3.4).

### 2.3 Preparación de los datos de entrenamiento

Todos los procedimientos se llevaron a cabo en una máquina virtual patrocinada por Microsoft Azure 2, con el sistema operativo Windows Server 2016 Datacenter, de tamaño estándar NC6 (6 vCPUs [Intel Xeon CPU E5-2690 v3 2,60 GHz], 56 GB de memoria RAM y un co-procesador GPU NVIDIA Tesla K80). Los algoritmos se implementaron utilizando el lenguaje de programación Python. Se utilizó el paquete Rasterio [20] para acceder y procesar los datos ráster geoespaciales, y Shapely [21] para la manipulación de polígonos.

Se utilizaron un total de tres (3) archivos ráster con las mismas dimensiones (ancho: 7942 px, alto: 2900 px) y tamaño de píxel (10 m) para ejecutar los clasificadores: la imagen de satélite de 10 bandas, un archivo raster con los datos de entrenamiento y un archivo raster con los datos de validación. Cabe destacar que la partición de los datos en los conjuntos de entrenamiento y validación se realizó a nivel de polígono lo cual permitió equilibrar los conjuntos de datos de entrenamiento y de validación y, al mantener juntos los píxeles vecinos se tuvieron en cuenta también los efectos de la textura [22]. Las dos capas raster correspondientes a los datos de entrenamiento y validación se generaron de la siguiente manera: los polígonos digitalizados a partir de los datos de referencia de campo se dividieron aleatoriamente en dos conjuntos de datos. Los polígonos se asignaron al grupo de entrenamiento hasta que la superficie acumulada (ha) alcanzó el 80% de la superficie total cubierta por los polígonos para esa clase (Tabla 2). Los polígonos restantes se asignaron al grupo de validación. Para garantizar que no hubiera polígonos únicos representando clases, los polígonos de gran tamaño se dividieron en otros más pequeños antes de la partición. Por último, los polígonos asignados a ambos grupos se convirtieron en archivos raster en los que los valores de los píxeles corresponden a una de las seis clases LULC definidas para este estudio. A los píxeles que no pertenecen a ninguna clase se les asignó un valor estándar -sin datos-. Cada píxel de la imagen de satélite

que representa una clase LULC en los conjuntos de datos de entrenamiento y validación se convirtió en un vector. Cada característica dentro del vector representa una banda espectral de la imagen de satélite, construyendo así una matriz  $M \times N$ , donde  $M$  es el número de píxeles etiquetados como pertenecientes a una clase de la imagen de entrenamiento y  $N$  el número de características (10 bandas/características en este caso). Esta matriz se utilizó como archivo de entrada para los métodos de aprendizaje automático utilizados.

**Table 2.** Partición de datos

			Cultivos Frutales	Horticultura	Arbustales	Agua	Construcciones	Pasturas
Entrenamiento (78.7% píxeles)	Tamaño de Imagen (247575, 10)	Poligonos	121	168	188	364	1358	431
		Píxeles	7130	18787	76853	37119	16791	89895
Validación (21.2% píxeles)	Tamaño de Imagen (66641, 10)	Poligonos	23	46	62	1	296	101
		Píxeles	1811	5050	21700	11084	4228	22780

## 2.4 Evaluación y comparación de clasificadores

Se seleccionaron siete métodos de aprendizaje automático supervisado utilizados para clasificación de imágenes de satélite para compararlos entre sí. Estos clasificadores fueron Nearest Neighbors (KNN), Random Forest, Multi-layer Perceptron, Adaptive Boosting (AdaBoost), Naive Bayes, Quadratic Discriminant Analysis (QDA) y Decision Tree [23], [24], [25]. Los diferentes algoritmos de clasificación se programaron con scikit-learn [26], que proporciona herramientas y funciones avanzadas de aprendizaje automático para calcular métricas de comparación. Cada clasificador se ajustó buscando el espacio de parámetros para la mejor puntuación de rendimiento. Para ello, se consideraron cuidadosamente todas las combinaciones de parámetros con una estrategia de búsqueda de parámetros eficiente. Los parámetros de cada clasificador utilizados para aplicar estas búsquedas se optimizan mediante una búsqueda cruzada en una cuadrícula de parámetros [27] Ver en [DOI: 10.5281/zenodo.5338597](https://doi.org/10.5281/zenodo.5338597).

Para evaluar el rendimiento global de cada clasificador y con el fin de compararlos, calculamos tres métricas típicas: La precisión de la clasificación que evalúa la proporción de píxeles en el conjunto de datos de prueba que se predijeron correctamente, el coeficiente Cohen's Kappa que expresa el nivel de acuerdo entre las predicciones y los datos de referencia, y la pérdida Hamming Loss para estimar los píxeles que se predijeron incorrectamente [9].

Durante el proceso de entrenamiento, la precisión y la pérdida se calcularon a partir de los datos de entrenamiento y de validación para evitar el sobreajuste y garantizar

un equilibrio entre el sobre-ajuste y el sub-ajuste. Esta representación del progreso del entrenamiento nos permitió observar una variación óptima del entrenamiento y la validación en cada época del proceso, y reflejó la estabilidad de los clasificadores.

Para evaluar el rendimiento de los modelos para cada clase de LULC se calcularon las matrices de confusión utilizando los valores de los píxeles de las clases de LULC predichas y del conjunto de datos de validación, que representa las clases de LULC observadas. También se evaluaron otras métricas estándar utilizadas para evaluar el rendimiento de los métodos de clasificación, como: Precisión, Recall y F-measure. En las métricas de clasificación, la Precisión es una métrica que cuantifica la calidad de predicción del modelo, es decir la proporción de datos (píxeles, en este caso) correctamente clasificados del total de datos etiquetados (datos de referencia, ver Sección 2.1). Una precisión de 1,0 para una clase significa que todos los elementos predichos como pertenecientes a esa clase habían sido asignados a esa clase. Mientras que Recall cuantifica cuántos de los elementos etiquetados fueron correctamente clasificados, un Recall de 1.0 indica que cada elemento etiquetado para una clase fue correctamente asignado a esa clase. F-measure (también llamada F1-score) es la media armónica de la Precisión y el Recall, es decir, una media ponderada de ambas [28]. Para tener en cuenta el desbalance de clases, las métricas anteriores se calcularon utilizando la media ponderada por el número de instancias verdaderas de cada clase.

### 3 Resultados

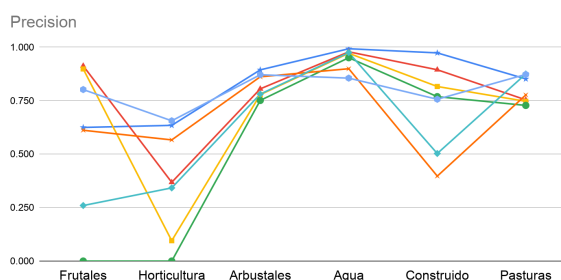
Los resultados analizados corresponden al mejor conjunto de parámetros seleccionados por cada clasificador. El tiempo de entrenamiento y validación de los algoritmos tradicionales osciló entre 1 y 2 horas. En la Tabla 3 se detallan los tiempos para cada algoritmo.

Nearest Neighbors mostró las mejores puntuaciones globales en comparación con los otros métodos y los peores rendimientos los obtuvieron AdaBoost y QDA (Tabla 3). Al examinar las puntuaciones de Precisión, Cohen's Kappa y Hamming Loss, Random Forest también tuvo un buen rendimiento con valores de precisión superiores al 80%. Sin embargo, cuando se analizó por clases, Random Forest tuvo problemas para reconocer las parcelas de Horticultura y de Cultivos Frutales (es decir, bajo valor de Recall) (Figura 2c).

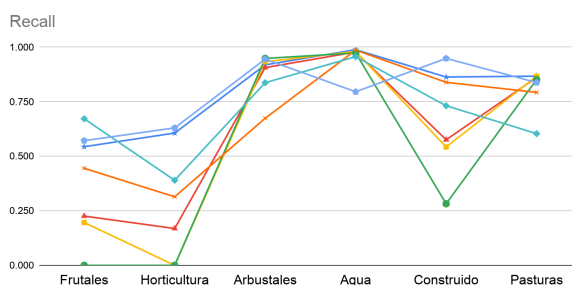
**Tabla 3.** Comparación de los métodos de clasificación.

	Nearest Neighbors	Decision Tree	Random Forest	Multi-layer Perceptron	AdaBoost	Naive Bayes	QDA
Precisión	0.889	0.832	0.814	0.794	0.773	0.748	0.725
Cohen's Kappa	0.849	0.764	0.736	0.707	0.696	0.670	0.636
Hamming Loss	0.111	0.168	0.186	0.206	0.227	0.252	0.275

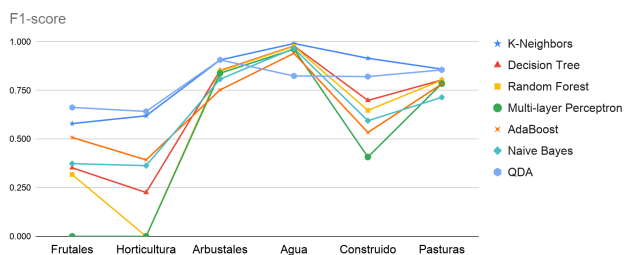
Tiempo de Entrenamiento	00:23	00:01	03:21	00:17	00:01	00:01	00:01
Tiempo de Validación	17:06	00:28	06:44	05:39	00:29	00:33	00:05



a)



b)



c)

**Fig. 2.** Comparación de los métodos por clase. (a) Métrica de precisión. (b) Recall y (c) F1-score.

Al analizar las métricas por clase, todos los clasificadores obtuvieron buenos resultados al clasificar las clases Agua, Arbustos y Pasturas, siendo KNN el que mostró las mejores puntuaciones de precisión (Figura 2a). Todos los clasificadores tuvieron dificultades para identificar los Cultivos Frutales. La capacidad de los



clasificadores para encontrar todas las muestras positivas se situó por debajo del 70% (Figura 2b). Los elevados valores de precisión para las clases Cultivos Frutales y Horticultura arrojados por Random Forest fueron engañosos porque este algoritmo sólo reconoció un número bajo de píxeles pertenecientes a esas clases (es decir, valores de Recall cercanos a 0, y en el caso de Decision Tree a 0.2). La métrica F-measure muestra que el modelo KNN superó a todos los clasificadores (Figura 2c), obteniendo los mejores valores en cuatro de las seis clases analizadas.

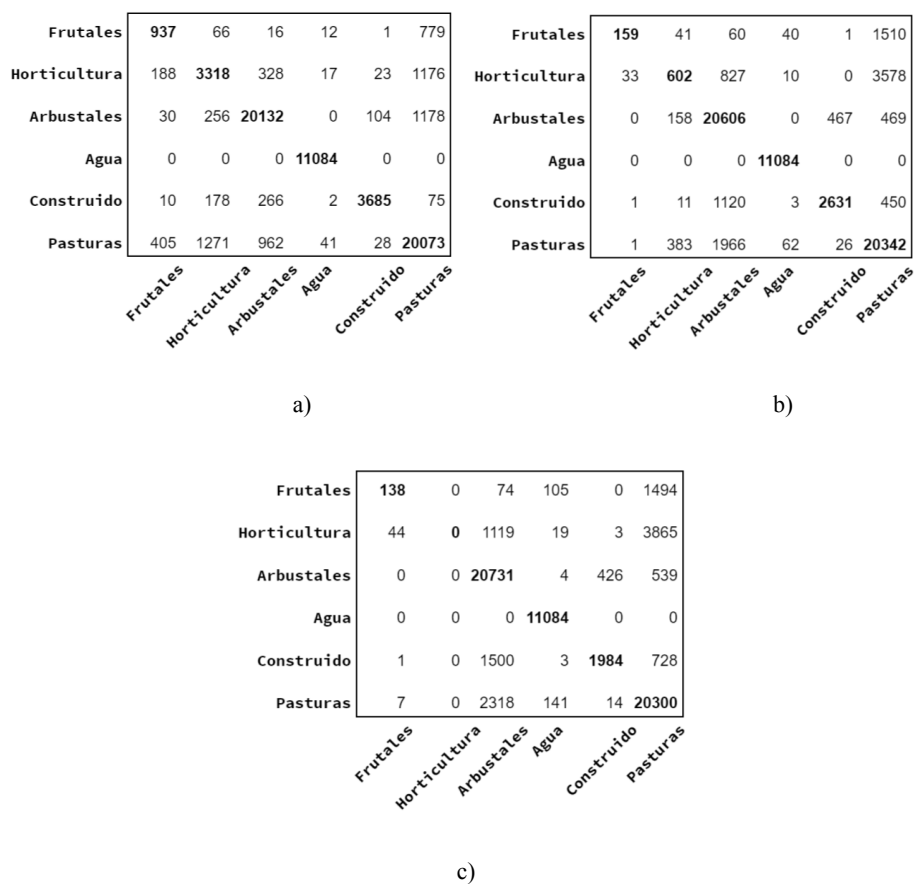


Fig. 3. Matrices de confusión para los clasificadores con mejores resultados. a) K-Nearest Neighbors, b) Decision Tree y c) Random Forest.

#### 4 Discusión y conclusiones

En este estudio hemos programado, entrenado y comparado 7 métodos diferentes de aprendizaje automático para la clasificación basada en píxeles de imágenes de satélite con el objetivo de generar un mapa de uso y cobertura del suelo, incluyendo cultivos, del valle agrícola situado en la Cuenca Inferior del Río Chubut. Aunque los mapas de LULC son insumos valiosos tanto para el sector académico, gubernamental y la sociedad civil este es el primer estudio que intenta automatizar y agilizar el proceso de generación de mapas para la zona. Este estudio permitió validar el uso de algoritmos de clasificación y valorar su desempeño para la construcción de mapas de LULC que sirvan de apoyo para el ordenamiento territorial y el delineamiento de políticas hídricas. Un aporte fundamental de este trabajo es la generación de código abierto que quedará disponible para mejorar la capacidad técnica, el manejo y la evaluación de algoritmos de clasificación en este u otros sistemas de estudio.

Con respecto al desempeño de los clasificadores evaluados los métodos K-Neighbors, Decision Tree y Random Forest, fueron los que presentaron los resultados más aceptables, con predicciones que superan el 80% de precisión global. Independientemente del método de clasificación, algunas clases fueron difíciles de discriminar. Ese fue el caso de Cultivos Frutales y Horticultura, que sistemáticamente se confundieron con Pasturas. Aunque incluimos las bandas espectrales de Sentinel-2 especialmente diseñadas para el monitoreo de la vegetación (es decir, las bandas red-edge), está claro que algunos aspectos técnicos deben mejorarse. Es muy probable que la confusión entre estas clases (Cultivos Frutales-Pasturas y Horticultura-Pasturas) se deba a las particularidades de los datos de entrenamiento. Nuestras áreas de entrenamiento estuvieron definidas por polígonos que encierran parcelas individuales que incluyen una clase principal de LULC (por ejemplo, Cultivos Frutales). Sin embargo, en el caso de los cultivos frutales, dentro de cada parcela individual los árboles frutales están plantados en filas y rodeados de pasto. Por lo tanto, cabe esperar que las parcelas en las que las pasturas gramíneas (pastos) son la principal cobertura del suelo se confundan a veces con las parcelas que contienen árboles frutales con una menor proporción de cobertura de pasto. En el caso de Horticultura, etiquetamos como "horticultura" a una variedad de cultivos que se cultivan en hileras en el campo (papa, cebolla, zanahoria). Por lo tanto, los píxeles etiquetados como "horticultura" abarcan una gran variedad de características ópticas. Probablemente se trate de limitaciones derivadas de trabajar únicamente con la respuesta espectral de las coberturas terrestres [29], [39]. Aunque hemos tenido en cuenta los efectos de vecindad al mantener juntos los píxeles de una determinada clase, incluir información sobre el contexto espacial puede mejorar los resultados de la clasificación al tener en cuenta la heterogeneidad espectral y estructural del píxel y su vecindad dentro de la misma clase de LULC. Incluir información sobre textura, por ejemplo, ha demostrado mejorar la caracterización de la estructura de la vegetación [30] y los resultados de la clasificación de LULC [29], [31], [39], [40]. Por otro lado, un par de cultivos que tienen una respuesta espectral muy similar la mayor parte del año, podrían diferenciarse en un momento concreto debido a las diferencias que

presentan en su ritmo de crecimiento y patrón fenológico. En este sentido, trabajar con imágenes multitemporales de Sentinel-2 que capturan no sólo la firma espectral de los cultivos sino también su firma temporal ha demostrado mejorar la clasificación de los mismos [32], [33]). En el caso de nuestro estudio, incluir imágenes adquiridas en diferentes fechas proporcionaría información valiosa sobre la fenología de los cultivos. Por ejemplo, los árboles frutales pierden sus hojas en otoño mientras que los cultivos hortícolas alcanzan su pico verde en verano. Aunque hemos realizado algunas pruebas utilizando imágenes multitemporales de Sentinel-2 y modelos de Deep learning con resultados preliminares prometedores, es necesario seguir trabajando para determinar las mejores fechas basándose en índices espectrales de vegetación y conocimiento local del manejo y variación temporal de los cultivos. Por otro lado, la incorporación de estos índices espectrales en la clasificación podría también mejorar los resultados como ya se ha observado en estudios anteriores [43], [44].

La calidad y la cantidad de los datos de entrenamiento son cruciales para la cartografía automatizada de LULC. En nuestro caso de estudio utilizamos datos de entrenamiento recogidos en el campo que eran representativos de seis clases principales de LULC. Se trata de un caso de estudio relativamente sencillo y es necesario realizar más pruebas para incluir más clases (por ejemplo, corrales de engorde, invernaderos, etc.) también presentes en este valle agrícola. Además, sólo hemos entrenado y probado los clasificadores con datos de campo recogidos a lo largo de un solo año.

De todos los algoritmos tradicionales de aprendizaje automático evaluados en este estudio, Random Forest (RF) es probablemente el más utilizado y reportado en la literatura. Se ha utilizado con éxito para la clasificación del uso/cobertura del suelo utilizando datos de teledetección en varios estudios [34]. Sin embargo, en nuestro estudio RF mostró un modesto rendimiento y no logró discriminar Cultivos frutales y Horticultura de Pasturas. Además de las limitaciones de nuestro caso de estudio mencionadas anteriormente, se ha comprobado que el rendimiento de RF mejora cuando se utilizan grandes conjuntos de datos de entrenamiento espectralmente homogéneos y equilibrados [34], [35], [36]. Como se ha mencionado anteriormente, nuestros datos de entrenamiento muestran una gran variabilidad espectral. Además, los cultivos frutales y la horticultura son los cultivos con menos presencia en el valle agrícola [41] y, por lo tanto, los menos representados en nuestro conjunto de datos de entrenamiento. En resumen, RF mostró menos capacidad para clasificar nuestro conjunto de datos, mientras que KNN mostró un mejor rendimiento general y por clase individual, demostrando que es un modelo más robusto.

Este trabajo fue desarrollado por un grupo interdisciplinario (profesionales con experticia en agricultura e irrigación, biología, ciencias de la computación, ingeniería electrónica, teledetección y meteorología) el cual continúa trabajando en el desarrollo de un modelo de clasificación basado en Deep Learning que mitigue las limitaciones obtenidas con los métodos tradicionales. El objetivo principal es proporcionar herramientas transferibles a las agencias gubernamentales y a los expertos en ciencias ambientales que podrían no tener experiencia específica en modelos de aprendizaje automático. Este estudio fue impulsado originalmente por el programa de Microsoft

IA para la Tierra (AI for Earth) que brindó asesoramiento, acceso a recursos de procesamiento y almacenamiento disponibles en la plataforma Azure.

## Referencias

1. Nur Hairunnisa Razaai, Saiful Arif Abdullah, and Muhammad Imam Hasan Reza. Identifying factors and predicting the future land-use change of protected area in the agricultural landscape of malaysian peninsula for conservation planning. *Remote Sensing Applications: Society and Environment*, 18:100298, 2020.
2. P Ruiz Empanza, N Hongkarnjanakul, D Rouquette, C Schwob, and L Mezeix. Land cover classification in thailand's eastern economic corridor (eec) using convolutional neural network on satellite images. *Remote Sensing Applications: Society and Environment*, 20:100394, 2020.
3. Heather Tallis, Taylor H Ricketts, Gretchen C Daily, and Stephen Polasky. *Natural capital: theory and practice of mapping ecosystem services*. Oxford University Press, 2011.
4. Ana L Liberoff, Silvia Flaherty, Pablo Hualde, Martín I García Asorey, Marilyn L Fogel, and Miguel A Pascual. Assessing land use and land cover influence on surface water quality using a parametric weighted distance function. *Limnologica*, 74:28–37, 2019.
5. M Terrado, V Acuña, D Ennaanay, H Tallis, and S Sabater. Impact of climate extremes on hydrological ecosystem services in a heavily humanized mediterranean basin. *Ecological indicators*, 37:199–209, 2014.
6. Ramita Manandhar, Inakwu OA Odeh, and Tiho Ancev. Improving the accuracy of land use and land cover classification of landsat data using post-classification enhancement. *Remote Sensing*, 1(3):330–344, 2009.
7. Shobitha Shetty. *Analysis of machine learning classifiers for lulc classification on google earth engine*. Master's thesis, University of Twente, 2019.
8. Aaron E Maxwell, Timothy A Warner, and Fang Fang. Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9):2784–2817, 2018.
9. James B Campbell and Randolph H Wynne. *Introduction to remote sensing*. Guilford Press, 2011.
10. Gerald Forkuor, Kangbeni Dimobe, Idriss Serme, and Jerome Ebagnerin Tondoh. Landsat-8 vs. sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in burkina faso. *GIScience & remote sensing*, 55(3):331–354, 2018.
11. Joel Segarra, Maria Luisa Buchailot, Jose Luis Araus, and Shawn C Kefauver. Remote sensing for precision agriculture: Sentinel-2 improved features and applications. *Agronomy*, 10(5):641, 2020.
12. Silvia Flaherty and M Pascual. Xii jornadas nacionales de geografía física “contribuciones de la geografía física a la gestión de los territorios” trelew–chubut–república argentina–11, 12 y 13 de abril de 2018.
13. Rose M Rustowicz. *Crop classification with multi-temporal satellite imagery*. Final Reports. Univ. Stanford, 2017.
14. Marc Wieland and Massimiliano Pittore. Performance evaluation of machine learning algorithms for urban pattern recognition from multi-spectral satellite images. *Remote Sensing*, 6(4):2912–2939, 2014.
15. ARS Marçal, JS Borges, JA Gomes, and JF Pinto Da Costa. Land cover update by supervised classification of segmented aster images. *International journal of remote sensing*, 26(7):1347–1362, 2005.
16. James R Carr. Spectral and textural classification of single and multiple band digital images. *Computers & Geosciences*, 22(8):849–865, 1996.

17. Ian L Thomas, Vivien M Benning, and Neville P Ching. Classification of remotely sensed images. 1987.
18. J Dungan. Spatial prediction of vegetation quantities using ground and image data. *International Journal of Remote Sensing*, 19(2):267–285, 1998.
19. QGIS Development Team et al. Qgis geographic information system. Open source geospatial foundation project, 2016.
20. Sean Gillies, B Ward, AS Petersen, et al. Rasterio: geospatial raster i/o for python programmers. URL <https://github.com/mapbox/rasterio>, 2013.
21. Sean Gillies, A Bierbaum, K Lautaportti, and O Tonnhofer. Shapely: manipulation and analysis of geometric objects. Available online: [github.com/Toblerity/Shapely](https://github.com/Toblerity/Shapely) (accessed on 15 June 2019), 2007.
22. Mihran Tuceryan and Anil K Jain. Texture analysis. *Handbook of pattern recognition and computer vision*, pages 235– 276, 1993.
23. Surekha Borra, Rohit Thanki, and Nilanjan Dey. *Satellite image analysis: clustering and classification*. Springer, 2019.
24. Jacinta Holloway and Kerrie Mengersen. Statistical machine learning methods and remote sensing for sustainable development goals: a review. *Remote Sensing*, 10(9): 1365, 2018.
25. Phan Thanh Noi and Martin Kappas. Comparison of random forest, k-nearest neighbor, and support vector machine classifiers for land cover classification using sentinel2 imagery. *Sensors*, 18(1):18, 2018.
26. Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *The Journal of machine Learning research*, 12:2825–2830, 2011.
27. Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media, 2009.
28. Yutaka Sasaki et al. The truth of the f-measure. 2007, 2007.
29. Chuanliang Sun, Yan Bian, Tao Zhou, and Jianjun Pan. Using of multi-source and multi-temporal remote sensing data improves crop-type mapping in the subtropical agriculture region. *Sensors*, 19(10):2401, 2019.
30. Eric M. Wood, Anna M. Pidgeon, Volker C. Radeloff, Nicholas S. Keuler, Image texture as a remotely sensed measure of vegetation structure, *Remote Sensing of Environment* 121 (2012) 516–526
31. S Jenicka. *Land Cover Classification of Remotely Sensed Images: A Textural Approach*. Springer Nature.
32. Zhiwei Yi, Li Jia, and Qiting Chen. Crop classification using multi-temporal sentinel-2 data in the shiyang river basin of China. *Remote Sensing*, 12(24):4052, 2020.
33. Francesco Vuolo, Martin Neuwirth, Markus Immitzer, Clement Atzberger, and Wai-Tim Ng. How much does multi-temporal sentinel-2 data improve crop type classification? *International journal of applied earth observation and geoinformation*, 72:122–130, 2018.
34. Mariana Belgiu and Lucian Dragut. Random forest in remote sensing: A review of applications and future directions. *ISPRS journal of photogrammetry and remote sensing*, 114:24–31, 2016.
35. Arun D Kulkarni and Barrett Lowe. *Random forest algorithm for land cover classification*. 2016.
36. Koreen Millard and Murray Richardson. On the importance of training data sample selection in random forest image classification: A case study in peatland ecosystem mapping. *Remote sensing*, 7(7):8489–8515, 2015.
37. Lucas D Díaz, Daniela A Raguileo, Marcos Hernández, and Franco J Salvadores. Caracterización del sistema de riego del valle inferior del río chubut: análisis desde las representaciones y opiniones de quienes riegan, 2021.

38. Ferran Gascon, Catherine Bouzinac, Olivier Thépaut, Mathieu Jung, Benjamin Francesconi, Jérôme Louis, Vincent Lonjou, Bruno Lafrance, Stéphane Massera, Angélique Gaudel-Vacaresse, Florie Languille, Bahjat Alhammoud, Françoise Viallefont, Bringfried Pflug, Jakub Bieniarz, Sébastien Clerc, Laëticia Pessiot, Thierry Trémas, Enrico Cadau, Roberto De Bonis, Claudia Isola, Philippe Martimort and Valérie Fernandez. (2017) Copernicus Sentinel-2A Calibration and Products Validation Status, *Remote Sensing*, 9, 584; doi:10.3390/rs9060584
39. Miao Li, Shuying Zang, Bing Zhang, Shanshan Li & Changshan Wu (2014), A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information, *European Journal of Remote Sensing*, 47:1, 389-411, DOI: 10.5721/EuJRS20144723
40. Kumar, R.S., Menaka, C. & Cutler, M.E.J. ANN Based Robust LULC Classification Technique Using Spectral, Texture and Elevation Data. *J Indian Soc Remote Sens* 41, 477–486 (2013). <https://doi.org/10.1007/s12524-012-0247-y>
41. HCA. 2013. Plan Director del Recursos Hídricos del Río Chubut. Informe Final. Abril de 2013. Provincia de Chubut. Consejo Federal de Inversiones. HCA Consultora S.R.L. 5 Tomos. <http://biblioteca.cfi.org.ar/documento/plan-director-de-recursos-hidricos-del-rio-chubutprovincia-del-chubut/>
42. Agencia Espacial Europea, 2015, Sentinel User Handbook and Exploitation Tools, Volumen 1, Revision 2
43. Le'an Qu, Zhenjie Chen, Manchun Li, Junjun Zhi and Huiming Wang, Accuracy Improvements to Pixel-Based and Object-Based LULC Classification with Auxiliary Datasets from Google Earth Engine, *Remote Sensing*. 2021, 13, 453. <https://doi.org/10.3390/rs13030453>
44. A. Karakacan Kuzucu and F. Bektas Balcik, Testing the potential of vegetation indices for land use/cover classification using high resolution data, *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, Volume IV-4/W4
45. Copernicus Sentinel data (2019). Recuperado de Copernicus Open Access Hub (Noviembre 2020), procesado por ESA.