

# Evaluación de algoritmos de clasificación en la plataforma *Google Earth Engine* para la identificación y detección de cambios de construcciones rurales y periurbanas a partir de imágenes de alta resolución

Alejandro Coca-Castro<sup>1</sup>\*, Maycol A. Zaraza-Aguilera<sup>1</sup>, Yilsey T. Benavides-Miranda<sup>1</sup>, Yeimy M. Montilla-Montilla<sup>1</sup>, Heidi B. Posada-Fandiño<sup>1</sup>, Angie L. Avendaño-Gómez, Hernando A. Hernández-Hamon, Sonia C. Garzón-Martínez<sup>1</sup>, Carlos A. Franco-Prieto<sup>1</sup>\*

<sup>1</sup> Grupo Investigación, Desarrollo e Innovación del Centro de Investigación y Desarrollo en Información Geográfica, Instituto Geográfico Agustín Codazzi, Carrera 30 N° 48-51, Bogotá D.C., Colombia..

---

**Resumen:** La detección de cambios de áreas construidas basada en datos de teledetección es una importante herramienta para el ordenamiento y la administración del territorio p.e.: la identificación de construcciones ilegales, la actualización de registros catastrales y la atención de desastres. Bajo el enfoque de comparación post-clasificación, la presente investigación tuvo como objetivo evaluar la funcionalidad de varios algoritmos de clasificación para identificar y capturar las construcciones y su cambio entre dos fechas de análisis usando imágenes de alta resolución (<1 m/píxel) en ámbitos rurales y límites del perímetro urbano municipal. La anterior evaluación fue llevada a cabo a través de una aplicación desarrollada mediante la plataforma *Google Earth Engine* (GEE), donde se alojaron y analizaron diferentes imágenes y datos de entrada sobre dos áreas de estudio en Colombia. En total, ocho algoritmos de clasificación tradicional, tres no supervisados (*K-means*, *X-Means* y *Cascade K-Means*) y cinco supervisados (*Random Forest*, *Support Vector Machine*, *Naive Bayes*, *GMO maximum Entropy* y *Minimum distance*) fueron entrenados empleando GEE. Adicionalmente, se entrenó una red neuronal profunda denominada *Feature Pyramid Networks* (FPN) sobre la cual se aplicó la estrategia de modelos preentrenados, usando pesos del modelo *EfficientNetB3*. Por cada una de las dos áreas de estudio, tres zonas de evaluación fueron propuestas para cuantificar la funcionalidad de los algoritmos mediante la métrica *Intersection over Union* (IoU). Esta métrica representa la evaluación de la superposición de dos regiones y tiene un rango de valores de 0 a 1, donde a mayor coincidencia de las imágenes mayor es el valor de IoU. Los resultados indican que los modelos configurados con la red FPN tienen la mejor funcionalidad, seguido de los algoritmos tradicionales supervisados. Las diferencias de la funcionalidad fueron específicas por área de estudio. Para el ámbito rural, la mejor configuración de FPN obtuvo un IoU promedio entre ambas fechas de 0,4, es decir, cuatro veces el mejor modelo supervisado, correspondiente al *Support Vector Machine* de kernel Lineal con un IoU de 0,1. Respecto al área de límites del perímetro urbano municipal, esta diferencia fue menos marcada, con un IoU promedio de 0,53 en comparación con el 0,38 derivado del mejor modelo de clasificación supervisada, que en

**To cite this article:** Coca-Castro, A., Zaraza-Aguilera, M.A., Benavides-Miranda, Y.T., Montilla-Montilla, Y.M., Posada-Fandiño, H.B., Avendaño-Gómez, A.L., Hernández-Hamon, H.A., Garzón-Martínez, S.C., Franco-Prieto, C.A. 2021. Evaluation of classification algorithms in the Google Earth Engine platform for the identification and change detection of rural and periurban buildings from very high-resolution images. *Revista de Teledetección*, 58, 71-88. <https://doi.org/10.4995/raet.2021.15026>

\* Corresponding author: [acocac@gmail.com](mailto:acocac@gmail.com), [carlos.franco@igac.gov.co](mailto:carlos.franco@igac.gov.co)

este caso fue *Random Forest*. Los resultados de esta investigación son relevantes para entidades responsables del seguimiento de las dinámicas de las *áreas construidas* a partir de plataformas de procesamiento en la nube como GEE, estableciendo una línea base para futuros estudios evaluando la funcionalidad de los clasificadores disponibles en otros contextos.

**Palabras clave:** clasificación, construcciones, catastro, computación en la nube, aprendizaje de máquinas, aprendizaje profundo.

## Evaluation of classification algorithms in the Google Earth Engine platform for the identification and change detection of rural and periurban buildings from very high-resolution images

**Abstract:** Building change detection based on remote sensing imagery is a key task for land management and planning e.g., detection of illegal settlements, updating land records and disaster response. Under the post-classification comparison approach, this research aimed to evaluate the feasibility of several classification algorithms to identify and capture buildings and their change between two time steps using very-high resolution images (<1 m/pixel) across rural areas and urban/rural perimeter boundaries. Through an App implemented on the Google Earth Engine (GEE) platform, we selected two study areas in Colombia with different images and input data. In total, eight traditional classification algorithms, three unsupervised (K-means, X-Means y Cascade K-Means) and five supervised (Random Forest, Support Vector Machine, Naive Bayes, GMO maximum Entropy and Minimum distance) available at GEE were trained. Additionally, a deep neural network named Feature Pyramid Networks (FPN) was added and trained using a pre-trained model, EfficientNetB3 model. Three evaluation zones per study area were proposed to quantify the performance of the algorithms through the Intersection over Union (IoU) metric. This metric, with a range between 0 and 1, represents the degree of overlapping between two regions, where the higher agreement the higher IoU values. The results indicate that the models configured with the FPN network have the best performance followed by the traditional supervised algorithms. The performance differences were specific to the study area. For the rural area, the best FPN configuration obtained an IoU averaged for both time steps of 0.4, being this four times higher than the best supervised model, Support Vector Machines using a linear kernel with an average IoU of 0.1. Regarding the setting of urban/rural perimeter boundaries, this difference was less marked, having an average IoU of 0.53 in comparison to 0.38 obtained by the best supervised classification model, in this case Random Forest. The results are relevant for institutions tracking the dynamics of building areas from cloud computing platforms like GEE and establish a baseline for future assessments of classifiers in likewise platforms in other contexts.

**Key words:** classification, buildings, cadastre, cloud computing, machine learning, deep learning.

---

## 1. Introducción

La observación de la superficie de la Tierra es una labor de demanda creciente, pues tiene múltiples usos como son: la vigilancia de los recursos naturales a nivel local, regional y nacional, el seguimiento de la cobertura terrestre, los cambios de uso de la tierra, o estudios ambientales, entre otros. Los datos de teledetección a partir de diferentes plataformas permiten una amplia y rica oferta de información a distintas resoluciones, siendo insumos para la detección de cambios. Este proceso se define como la identificación de variaciones en el estado de un objeto o fenómeno sobre la superficie de la Tierra producidos a lo largo de un período de tiempo (Han et al., 2020).

La detección de cambios a partir de datos de teledetección cuenta con una diversidad de técnicas

como las que analizan diferencias en distintos niveles como píxel, objeto o extracción de patrones (Hussain et al., 2013). Los estudios con respecto a la detección de cambios en construcciones pueden dividirse en dos enfoques principales (Liu et al., 2019): a) los métodos soportados en la comparación post-clasificación; y b) los métodos basados en la detección directa. Los primeros extraen a partir de imágenes las construcciones de una zona de estudio en dos fechas distintas, para luego realizar la comparación entre los mapas de clasificación resultantes. Para el caso de los ámbitos rurales, que se caracterizan generalmente por construcciones pequeñas y dispersas, este tipo de enfoque se ha realizado con éxito mediante arquitecturas *deep learning*. Por ejemplo, Vargas-Muñoz et al. (2019) propusieron una red convolucional (CNN) capaz de predecir construcciones a partir de datos abiertos

de *Open Street Map* (OSM) e imágenes de 0,60 m/píxel en ámbitos rurales de Tanzania y Zimbabue. Los autores resaltaron que las salidas de la red combinadas con un post-procesamiento basado en los campos aleatorios de Markov mejoran substancialmente la delineación y localización de las construcciones en este tipo de ámbitos. He et al. (2019), por su parte, desarrollaron una red *multiscale siamese* para la identificación de construcciones tanto en entornos urbanos como rurales a partir de imágenes de drones e información de altura derivada de estas en la provincia de Jiangxi, China. La ventaja de la red es que no requiere de muestras de entrenamiento y logra extraer eficientemente las características finas y gruesas de las construcciones a partir de las imágenes de entrada. Adicionalmente, en comparación con otras redes como *U-Net* y *Fully Connected Network* (FCN), la red propuesta consume menos recursos hardware y por ende requiere un menor tiempo de entrenamiento. La capacidad de generalización de los modelos *deep learning* permite que estos traspasen el campo experimental, encontrándose implementaciones para la identificación de construcciones en áreas relativamente grandes. Por ejemplo, Yang et al. (2018) lograron generar predicciones de las áreas construidas para toda la superficie de los Estados Unidos a partir de la evaluación de varios modelos de redes convolucionales entrenados con imágenes de cuatro bandas espectrales (RGB e infrarrojo) de 1 m/píxel de resolución espacial y una cantidad considerable de muestras de entrenamiento distribuidas sobre todo el territorio.

Con relación al conjunto de métodos de detección directa, estos se basan en un marco integral de análisis en el que se han entrenado modelos que permiten hacer la tarea en un solo flujo de trabajo, sin necesidad de realizar la clasificación previa de las imágenes. Este enfoque, por su complejidad, resulta por el momento limitado al ámbito investigativo y mayormente aplicado en áreas urbanas. Una propuesta diversa de modelos *deep learning* se encuentran en la literatura como redes siamesas (Daudt et al., 2018; Liu et al., 2019) y *Mask R-CNN Fusion Sobel* (Zhang et al., 2020), de las cuales se ha reportado una mayor funcionalidad respecto a los algoritmos tradicionales (Ma et al., 2016).

El procesamiento en la nube ha tenido una alta acogida para evaluar diferentes estudios

en teledetección, entre ellos, los asociados con las técnicas de detección de cambios en construcciones gracias a la posibilidad del manejo y análisis de grandes volúmenes de datos. En este sentido, la compañía Google con su producto *Earth Engine* (GEE) ofrece un catálogo con petabytes de imágenes satelitales y conjuntos de datos geoespaciales con capacidades de análisis a escala planetaria que están a disposición de científicos, investigadores y desarrolladores. Una mayor información de las bondades y usos de esta plataforma se encuentra en Gorelick *et al.* (2017).

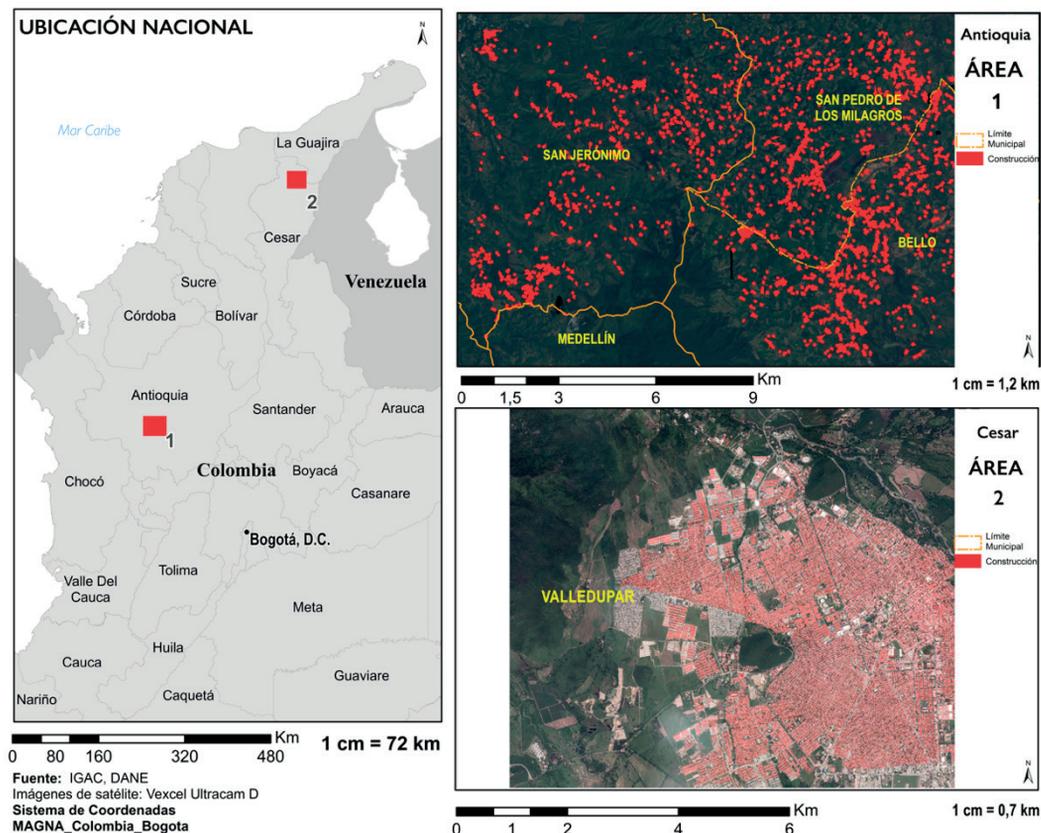
En términos prácticos, las metodologías de detección de cambios y plataformas como GEE resultan de interés para su implementación en procesos de formación y actualización catastral. El primero se refiere a la obtención de información sobre los terrenos y construcciones correspondientes a los predios de una unidad orgánica catastral o parte de ella con el fin de alimentar el censo catastral en los aspectos físico, jurídico y económico. El segundo recopila información de los componentes físico y jurídico para actualizar los datos de la formación. Bajo este contexto, la presente investigación surge de la necesidad manifiesta de la Subdirección de Catastro del Instituto Geográfico Agustín Codazzi (IGAC) de mejorar estos procesos buscando el uso de nuevas tecnologías que permitan mayor eficiencia y fortalecimiento de la política pública de Catastro Multipropósito en Colombia. Considerando que los procesos catastrales mencionados son complejos y costosos debido al uso de técnicas tradicionales, se demanda el desarrollo de técnicas innovadoras para la identificación y captura de los objetos de interés

De acuerdo con la revisión de la literatura y las necesidades planteadas, el presente trabajo tiene por objeto la identificación de las construcciones y la detección de su cambio a través del tiempo en ámbitos rurales y límites del perímetro urbano municipal usando GEE.

## 2. Áreas de estudio y Materiales

### 2.1. Áreas de estudio

Las áreas de estudio contemplaron dos sectores del país con cobertura de imágenes de alta resolución



**Figura 1.** Ubicación de las áreas de estudio. Los vectores de color rojo sobrepuestos sobre las imágenes satelitales indican las construcciones obtenidas por Catastro.

espacial, pero de proveedores diferentes (Figura 1). Esto permitió la comparación de los métodos de detección de cambios en construcciones en condiciones diferentes. El área 1, que representa un ámbito rural, se encuentra ubicada en el departamento de Antioquia, en la zona denominada Valle de Aburrá, con una superficie de 150 km<sup>2</sup>. Se caracteriza por tener una topografía irregular con alturas entre 1300 y 2800 m s.n.m. y temperatura promedio de 17,33 °C (Vásquez *et al.*, 2006). El área 2, que representa un ámbito urbano, se encuentra en el departamento del César, en la zona de la cabecera municipal de Valledupar y parte rural colindante a esta cabecera. El municipio está ubicado al norte del Valle del César, al margen de los ríos César y Guatapurí. Este tiene una superficie de 123 km<sup>2</sup>, presenta una temperatura media de 28,4 °C y altitud entre los 150 y 220 m s.n.m. (Pinzón-Rodríguez y Sanabria-García, 2017).

## 2.2. Insumos de entrada

### 2.2.1. Imágenes de alta resolución

La selección de las imágenes de alta resolución fue realizada considerando las especificaciones técnicas para productos de información en procesos de formación y actualización catastral con enfoque multipropósito dados por el IGAC (2020), especialmente con las resoluciones espaciales para ortoimágenes. Dos tipos de insumos fueron manejados en esta investigación: imágenes satelitales y aerofotografías. La imagen satelital en el área de Antioquia correspondió a la misión satelital PeruSAT-1 con año de toma 2017, mientras para el área de Valledupar se empleó WorldView-2 del año 2014. La Tabla 1 presenta las características de las imágenes satelitales empleadas. Respecto a las aerofotografías, estas correspondieron al sensor Vexcel Ultracam D, las cuales en su toma original presentaron una distancia de muestra del suelo (GSD) de 0,15 m;

**Tabla 1.** Especificaciones de las imágenes satelitales y sus resoluciones.

Área de estudio	Insumo	Fecha captura	Resolución			
			Espacial	Espectral	Radiométrica	Temporal
Antioquia	PeruSAT-1	01/09/2017	Pancromática (0,7 m)	Pancromática (451-745 nm)	12 bits	26 días
			Multiespectral (2,8 m)	Azul (454-525 nm) Verde (533-598 nm) Rojo (633-702 nm) Infrarrojo cercano (764-884 nm)		
Valledupar	WorldView-2	23/12/2014	Pancromática (0,46 m)	Pancromática (450-800 nm)	16 bits	1,1 a 3,7 días
			Multiespectral (1,85 m)	Azul (450-510 nm) Verde (510-580 nm) Rojo (630-690 nm) Infrarrojo cercano (770-895 nm)		

**Tabla 2.** Especificaciones de las fotografías aéreas.

Área de estudio	Sensor	Fecha(s) de captura	GSD (m)	Cantidad fotografías	Resolución espectral
Antioquia	Vexcel Ultracam D	20/03/2010	0,15	594	Pancromática (430-660 nm)
		21/03/2010			Azul (430-510 nm)
		23/03/2010			Verde (490-5800 nm)
Valledupar	Vexcel Ultracam D	23/10/2019	0,15	175	Rojo (580-670 nm) Infrarrojo cercano (700-880 nm)

no obstante, el mosaico fue generado con un GSD de 0,50 m. La Tabla 2 dispone las características de las aerofotografías.

### 2.2.2. Vectores históricos de construcciones

Se emplearon vectores de construcciones de 2010 a 2020 y 2012 a 2020 para las áreas de Antioquia y Valledupar, respectivamente, para la definición de muestras y evaluación de los resultados de clasificación de los algoritmos. Para efectos de la detección de cambio, se filtraron los vectores que corresponden a cada una de las fechas de las imágenes analizadas para cada área de estudio. La información espacial fue suministrada por los gestores catastrales oficiales, en este caso el IGAC y la Gobernación de Antioquia.

## 3. Metodología

La Figura 2 presenta el esquema de las fases de desarrollo consideradas para alcanzar los objetivos de este proyecto, denominadas así: selección de las áreas de estudio (Sección 2.1); colección y preparación de los insumos de entrada (Sección 2.2); validación de la correspondencia (Sección 3.1); ejecución de los algoritmos de clasificación (Sección 3.2), verificación y

evaluación de exactitud temática (Sección 3.3); y detección de cambio y vectorización (Sección 3.4).

### 3.1. Validación de correspondencia

En esta etapa se realizó la evaluación de los vectores históricos de construcciones con respecto a las imágenes de ambas áreas de estudio. Este procedimiento permitió depurar los vectores a ser involucrados en los procesos de muestreo y aprendizaje en las técnicas de clasificación supervisada. Adicionalmente, los vectores depurados permitieron la definición de tres zonas de evaluación, cada una con una superficie de 0,25 km<sup>2</sup>, donde se evidencian cambios para cada área de estudio. La validación de correspondencia consistió en identificar inconvenientes de topología y agrupar los polígonos de acuerdo con su exactitud y presencia de nubes. Las inconsistencias de topología fueron evaluadas con el cálculo del índice de compacidad (Ecuación 1) e índice de forma (Ecuación 2).

$$C = \frac{4 \pi \text{área}}{\text{perímetro}^2} \quad (1)$$

$$SI = \frac{\text{perímetro}}{4 \sqrt{\text{área}}} \quad (2)$$



**Figura 2.** Flujo de trabajo.

Para un objeto de forma circular el índice de compacidad es igual a 1, por lo que todas aquellas construcciones donde su compacidad fue cercana a cero no se tuvieron en cuenta ya que estas construcciones tienen una forma alargada. Por otra parte, el índice de forma para un polígono regular es igual a uno, que correspondería con un cuadrado exacto, siendo esta una forma muy inusual con respecto a la de una construcción.

De acuerdo con lo anterior, para el área de estudio 1, de un total de 11 259 polígonos de construcción, quedaron 3192 agrupados en cuatro categorías: cambio, nube, baja exactitud, alta exactitud. Para el área de estudio 2, de un total de 95 130 polígonos iniciales, se obtuvieron 75 245 agrupados en cuatro categorías: cambio, baja exactitud, alta exactitud, sin capturar.

### 3.2. Clasificación

La identificación y detección de cambios se basó en el método de comparación post-clasificación. En este componente se evaluaron diferentes técnicas tradicionales de clasificación supervisada y no supervisada (ver Sección 3.2.1), así como supervisadas usando redes neuronales profundas o *deep learning* (ver Sección 3.2.2).

#### 3.2.1. Algoritmos tradicionales

La versión usada de la plataforma GEE provee cerca de nueve tipos de algoritmos de métodos

tradicionales supervisados y cinco no supervisados. Mediante una evaluación de prefactibilidad se revisaron varios de estos y aquellos con mayor correspondencia para la identificación y detección de cambios se muestran en la Tabla 3.

Como datos de entrada, los clasificadores tradicionales usaron las cuatro bandas (azul, verde, rojo e infrarrojo) de los insumos de imágenes de alta resolución disponibles, esto para cada área de estudio. Adicionalmente, fue agregado el índice de vegetación normalizado (NDVI), el cual resalta las cubiertas vegetales facilitando la discriminación de construcciones.

En lo que se refiere a la colección de muestras para entrenamiento, se tuvo un tratamiento diferencial entre los algoritmos no supervisados y supervisados. Para el primer grupo, de acuerdo con el tamaño manejado en las zonas de evaluación, se definió que cada uno de los clúster o clases definidas por el algoritmo no supervisado tuviera un tamaño mínimo de 10 píxeles para la definición de los subgrupos de muestras. Para los clasificadores supervisados, el tamaño de muestra fue de 50 puntos por cada clase (área construida y área no construida), para un total de 100 por cada zona de evaluación. Estas muestras fueron adquiridas teniendo como referencia las capas depuradas de construcciones por cada imagen por área de estudio, manejando una partición de 70% y 30% para entrenamiento y evaluación.

**Tabla 3.** Lista de algoritmos de clasificación tradicional con descripción de su tipo, familia y configuraciones (hiperparámetros) usados en esta investigación. N/A se refiere a que no aplica.

Algoritmo	Tipo	Configuraciones
K-Means	No supervisado	<i>Distancia</i> {euclidiana, manhattan}, <i>número de clusters</i> {4}
X-Means	No supervisado	<i>Distancia</i> {euclidiana, manhattan, chebyshev}, <i>rango búsqueda de clusters</i> {2 a 6}
Cascade KMeans	No supervisado	<i>Distancia</i> {euclidiana, manhattan}, <i>rango búsqueda de clusters</i> {2 a 6}
Random Forest	Supervisado	<i>Número de árboles</i> {100}
Support Vector Machine	Supervisado	<i>Kernel</i> {lineal, polinómico, radial y sigmoidal}
Naive Bayes	Supervisado	N/A
GMO máximo Entropy	Supervisado	N/A
Minimum distance	Supervisado	<i>Distancia</i> {euclídea, cosine, mahalnobis}

### 3.2.2. Algoritmos deep learning

Teniendo en cuenta el particular éxito de las redes neuronales profundas para la identificación y detección de cambios en construcciones, estas fueron evaluadas en la presente investigación. Específicamente, las arquitecturas seleccionadas de aprendizaje supervisado *deep learning* se basan en redes neuronales convolucionales (CNN por sus siglas en inglés). De las diferentes variantes que existen de esta red, se emplearon aquellas óptimas para la tarea de segmentación semántica. El uso de esta tarea para aplicaciones de teledetección se compone secuencialmente de subredes de codificador y decodificador.

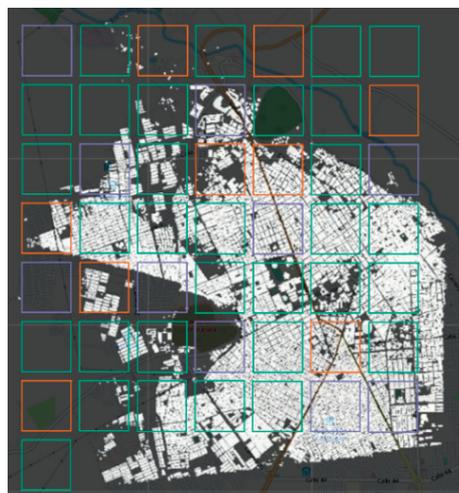
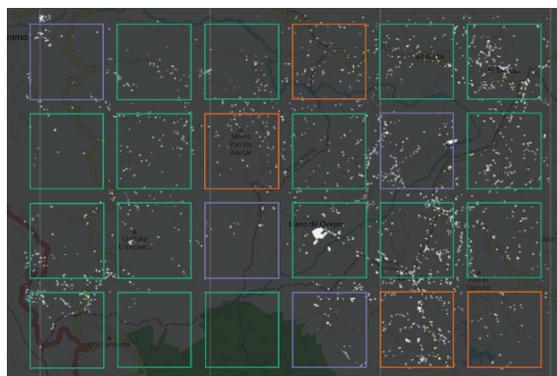
Los datos de entrada correspondieron a las bandas espectrales azul, verde y rojo obtenidas de las imágenes disponibles para cada área de estudio. La decisión de usar estas bandas y excluir la banda del infrarrojo y el índice de vegetación corresponde a la estrategia de entrenamiento con modelos preentrenados. Esta estrategia se refiere a acelerar la extracción de atributos de la imagen de interés a través de las características aprendidas por modelos que fueron entrenados usando millones de imágenes RGB, generalmente a partir del conjunto de datos de ImageNet (Li et al., 2009). Para una extracción exitosa de las características aprendidas a otras imágenes, es ideal que estas conserven la misma configuración de bandas. Aunque la banda de infrarrojo se podría incorporar de manera adicional, no todas las imágenes usadas para las actualizaciones catastrales tienen esta banda, por ejemplo, las

aerofotografías. Adicionalmente, tal como fue planteado por Yang et al. (2018), incorporar la banda de infrarrojo implica un procesamiento adicional para integrarse satisfactoriamente con el uso de modelos preentrenados. A diferencia de los modelos tradicionales, los modelos CNN usualmente requieren que los datos de entrada estén como *parche*. Este arreglo se refiere a no sólo tomar un píxel central sino su contexto en cada una de las bandas seleccionadas. De esta manera los datos de entrada se pueden formular como:

$$P(i)=X \times Y \times B$$

Donde un parche de entrada  $P(i)$  se conforma por unas dimensiones (número de píxeles) del eje  $X$  e  $Y$ , establecidas según el modelo y aplicación, y de  $B$  que se refiere a bandas  $B$  o canales de información seleccionados.

Debido a que la toma de muestras tiene en cuenta el contexto de píxel o parche, es recomendado que la estrategia de partición tenga un componente espacial. Siguiendo este lineamiento, que se ha aplicado en investigaciones anteriores, por ejemplo, Rußwurm y Körner (2018), se creó una grilla de bloques sobre cada área de estudio. El tamaño de estos bloques fue definido como 2,5 km×2,5 km con un margen de 0,2 km para Antioquia, y 0,85 km×0,85 km con un margen de 0,064 km para Valledupar. El margen permitió que los parches de cada partición no se superpusieran espacialmente, evitando sobreajustes de los modelos (Figura 3).



Bloques: □ Entrenamiento, □ Validación, □ Evaluación.

**Figura 3.** Particiones con el enfoque de bloques bajo una proporción 4:1:1 en Antioquia (izquierda) y Valledupar (derecha). El mapa base es *Open Street Map* y se dispone la capa de los vectores de construcciones disponibles (color blanco) por zona.

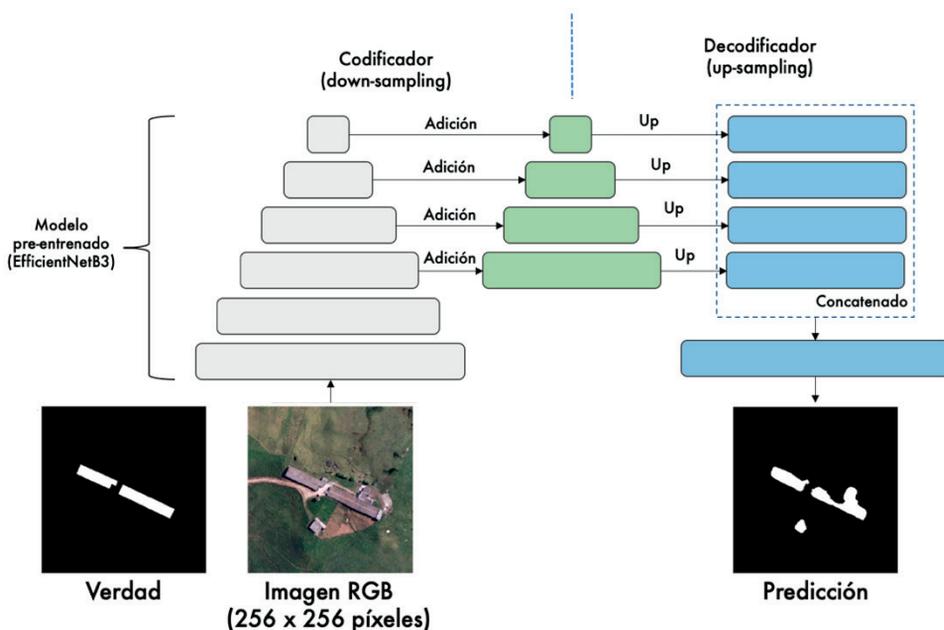
Posterior a la definición de estos bloques, se crearon puntos aleatorios sobre cada uno de ellos, los cuales correspondieron a 250 por clase. Sobre estos puntos se generaron los parches necesarios para entrenar los modelos de *deep learning* usando un tamaño de kernel de 256 píxeles×256 píxeles. Este valor fue escogido ya que es el comúnmente usado para entrenar los modelos de segmentación semántica evaluados. Cabe indicar que no todos los bloques contaban con el número objetivo de 250 muestras por clase, por lo que hubo ciertos bloques que solo tenían parches que pertenecían a una clase. El número de parches por área de estudio se presentan en la Tabla 4.

Para esta investigación se implementó la arquitectura de segmentación llamada *Feature Pyramid Networks* (FPN). Las FPN fueron propuestas para poder manejar imágenes y objetos expuestos a variaciones de escalas. La Figura 4 presenta la estructura de la red FPN implementada. La arquitectura se compone de capas multiescala

producidas a partir de una estructura piramidal de abajo-arriba (*bottom-up*), de arriba-abajo (*top-down*) y por el uso de conexiones laterales. Adicionalmente, debido a su enfoque de capas multiescala, las FPN permiten la adición de redes preentrenadas o *backbone* en la estructura abajo-arriba. Esto último favorece los tiempos de entrenamiento reduciendo considerablemente el consumo de recursos. El uso de las FPN ha sido exitoso en estudios de teledetección de acuerdo con un trabajo previo en imágenes de drones (Parmar et al., 2020) con el objetivo de clasificar seis clases, entre ellas construcciones. En este estudio, los autores evaluaron cuatro arquitecturas principales para la segmentación de imágenes de drones entre ellas: U-Net (Ronneberger et al., 2015), Linknet (Chaurasia y Culurciello, 2018), *Pyramid Scene Parsing Network* (PSPNet) (Zhao et al., 2017) y *Feature Pyramid Networks* (FPN) (Lin et al., 2017). En conjunto sobre estas arquitecturas, los autores evaluaron el uso de modelos preentrenados. Específicamente,

**Tabla 4.** Número de parches por partición por tiempo (imagen inicial / imagen final) por área de estudio. El número de píxeles en cada parche es equivalente a 65.536 (256 píxeles×256 píxeles).

Área de estudio	Entrenamiento	Validación	Evaluación	Total
Antioquia	1500 / 1550	400 / 400	400 / 400	2300 / 2350
Valledupar	2800 / 2800	750 / 750	900 / 900	4450 / 4550



**Figura 4.** Esquema de la arquitectura FPN. La estructura de abajo-arriba está compuesta por una serie de convoluciones. La imagen de entrada pasa por una capa convolucional y el mapa de características obtenido forma el primer nivel de la pirámide. En total, cinco niveles se forman con la cadena de convoluciones. El camino de arriba-abajo aumenta la resolución de los mapas de características. La resolución es aumentada en un factor de 2 en cada nivel usando el algoritmo *nearest neighbour*. Esto permite la propagación de la información de alto nivel por toda la pirámide. Las conexiones laterales (convoluciones  $1 \times 1$ ) ayudan a la reconstrucción de las relaciones espaciales de la imagen original.

se encontró que el modelo preentrenado *EfficientNetB3* provee una buena funcionalidad respecto a los otros dos posibles, *MobileNetv2* e *InceptionResNetv2*, con un número reducido de operaciones de punto flotante (FLOPS). Partiendo de los hallazgos de esta investigación, el modelo *EfficientNetB3* fue usado en la estructura de arriba-abajo para acelerar la extracción de características de las imágenes por cada fecha y área de estudio.

Como es común en el entrenamiento de modelos *deep learning*, adicional a la selección de la arquitectura y estrategia de inicio de los pesos, un componente importante es definir el tamaño

del lote, el optimizador y el número de épocas. El tamaño de lote definido fue de 16 con un optimizador *Adam* (Kingma y Ba, 2014). Además de entrenar modelos con los datos disponibles por área de estudio, se probaron como estrategias complementarias entrenar los modelos combinando los datos de ambas fechas en cada área de estudio y otra estrategia combinó todos los datos disponibles en ambas áreas y fechas. Esto se hizo para probar la utilidad de estas combinaciones en la predicción de una o ambas áreas de estudio. La Tabla 5 indica el número de combinaciones posibles entrenadas para este proyecto con el

**Tabla 5.** Diferentes combinaciones y número de épocas consideradas usando un tamaño de lote de 16.

Experimento	Datos de entrada	Número total de parches de entrenamiento	Número de épocas
Antioquia	Perusat-1 ó Vexcel (individual)	1500 / 1550	50
	Perusat-1 & Vexcel (combinado)	3050	25
Valledupar	Worldview-2 ó Vexcel (individual)	2800 / 2800	50
	Worldview-2 & Vexcel (combinado)	4600	25
Global	Perusat-1 & Worldview-2 & Vexcel Antioquia & Vexcel Valledupar	7650	15

número correspondiente de épocas establecido empíricamente de acuerdo con la cantidad de datos de entrenamiento.

En lo referente a la función de costo, esta cuantifica el nivel de un modelo para desarrollar de forma óptima la tarea asignada. La formulación de dicha función varía según el problema. Para tareas de clasificación como segmentación semántica existen varios tipos de funciones de costo. Para esta investigación fue usada la función de Focal-Tversky (Abraham y Khan, 2019), la cual está diseñada para reducir el desbalance de clases, entendido como un mayor número de muestras de áreas no construidas que construidas, siendo este aspecto dominante en Antioquia. Junto a la función de costo, para este trabajo se usó la métrica de *Intersection over union* (IoU) (Jaccard, 1908) que permitió determinar el modelo óptimo asociado con cierto número de épocas. Como se profundiza en la Sección 3.3, IoU mide el coeficiente de similitud de la clase objetivo entre dos áreas, siendo común su reporte en los estudios sobre la identificación de construcciones.

### 3.3. Post-procesamiento

Como resultado de cada algoritmo descrito en la Sección 3.2 se obtuvieron imágenes clasificadas con dos clases: área construida y área no construida; no obstante, la separabilidad espectral entre la cobertura de construcciones y suelo desnudo puede ser baja, en especial en los algoritmos tradicionales. Para reducir dichos efectos o clasificaciones erradas, se emplearon capas vectoriales de vías que fueron suministradas por la Subdirección de Cartografía. Los vectores de vías se transformaron a formato ráster con un *buffer* promedio de 2,5 m alrededor del objeto. Las capas binarias fueron usadas para enmascarar la salida inicial del clasificador, filtrando los valores que se catalogaron erróneamente como construcción siendo una vía. Este procedimiento garantizó una reducción del error en la clasificación y mejora significativa de la identificación de la capa de construcciones.

### 3.4. Evaluación cuantitativa

Este componente fue realizado mediante la evaluación de métricas derivadas de la matriz de confusión, así como la métrica de *Intersection*

*over Union* (IoU). Como su nombre lo indica, la métrica IoU es el resultado de la intersección sobre el criterio de unión y va de un rango de 0 a 1, donde 1 indica una alta correspondencia entre la predicción y el área de referencia. Las métricas fueron calculadas considerando los vectores depurados de construcciones disponibles en las zonas de evaluación en cada área de estudio y fecha. La métrica IoU fue usada como criterio principal para calificar los mejores algoritmos en la tarea de identificación de construcciones y a su vez en la detección de cambios.

### 3.5. Detección del cambio y vectorización

Posterior a la limpieza de las clasificaciones de cada uno de los algoritmos generada con la capa de vías, se confrontaron los mapas post-procesados de ambas fechas con el fin de detectar el cambio. Esta operación se hizo mediante la diferencia entre los píxeles de la clase área construida para las dos fechas analizadas. Tras la detección del cambio, la capa ráster fue convertida a formato vectorial. Este procedimiento se llevó a cabo tanto en la capa de construcciones de cada fecha como en la correspondiente al cambio. A todos estos vectores se les hizo un cálculo en términos de área en unidades de hectárea.

### 3.6. Aplicación y software

A través de GEE Apps se generó una aplicación que permitió: cargar y visualizar los insumos de entrada; entrenar los algoritmos tradicionales; hacer una predicción y evaluación cuantitativa, y realizar la detección de cambios y vectorización. La aplicación aceleró la experimentación y generación de resultados de acuerdo con las diferentes configuraciones y zonas de evaluación de esta investigación.

Los principales procedimientos que se hicieron fuera de GEE fueron la preparación de insumos (mosaicos), correspondencia de los vectores de construcciones, y el entrenamiento de los modelos *deep learning*. La preparación de las imágenes WorldView-2 y PeruSAT-1 se llevó a cabo en el software PCI-Geomatica versión 2016 y Banff 2019, respectivamente. Las aerofotografías Vexcel se procesaron en PCI-Geomatica versión 2018. En ArcMap versión 10.8 se realizó la verificación de

la correspondencia entre vectores e imágenes. El entrenamiento de los modelos de *deep learning* se hizo mediante notebooks de Colab, los cuales ofrecen acceso gratuito a unidades gráficas de procesamiento (GPU, por sus siglas en inglés), en este caso Tesla K80. La predicción de los modelos se logró mediante la integración de Google AI y GEE. Finalmente, la generación de gráficos se hizo con la librería Seaborn v.0.11 en Python versión 3.6.

#### 4. Resultados y discusión

De acuerdo con la métrica IoU, se identificó que aquellos modelos entrenados de *deep learning* supervisado, específicamente la red FPN, fueron superiores respecto a los tradicionales, considerando los resultados para ambas áreas de estudio en las tres zonas de evaluación preseleccionadas (donde se evidenció cambio) (Figura 5). Esta superioridad es más notable en el área de estudio de ámbito rural en Antioquia respecto al área de límites del perímetro urbano municipal en Valledupar. Los valores altos de

IoU en la zona 3 de Valledupar se relacionan con la alta densidad de construcciones, siendo estas más fáciles de capturar por la mayoría de los algoritmos de clasificación en comparación con las otras zonas. Los resultados concuerdan con investigaciones existentes comparando redes neuronales de segmentación semántica y algoritmos tradicionales de clasificación para la identificación de construcciones en ámbitos urbanos (Han et al., 2020) y/o rurales (He et al., 2019). En términos de la superficie de área construida, se observó una baja funcionalidad en los algoritmos tradicionales no supervisados, los cuales la sobrestimaron generando errores en la detección y cálculo del área de cambio (Figura 6). En contraste, los modelos *deep learning* tienen unas estimaciones más realistas y consistentes, con áreas de cambio no superiores a 1 y 3 hectáreas en Antioquia, y Valledupar, respectivamente (Figura 7).

Como parte del análisis, resultó importante estudiar de manera desagregada los mejores algoritmos por cada familia. De esta manera, la

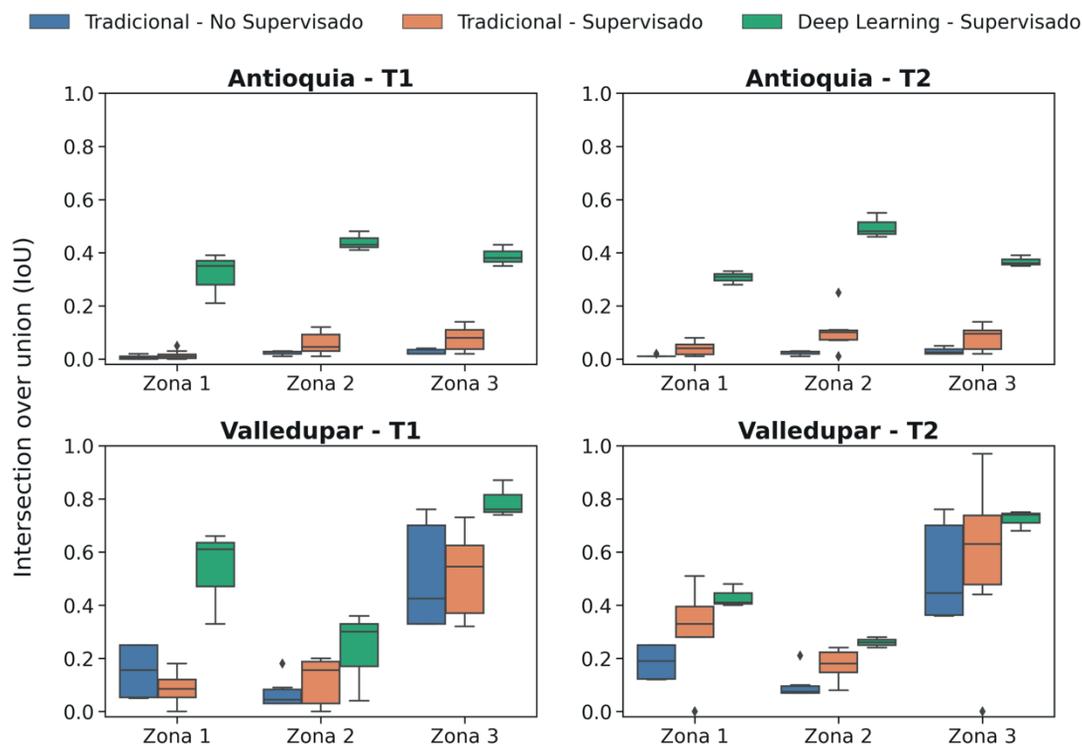
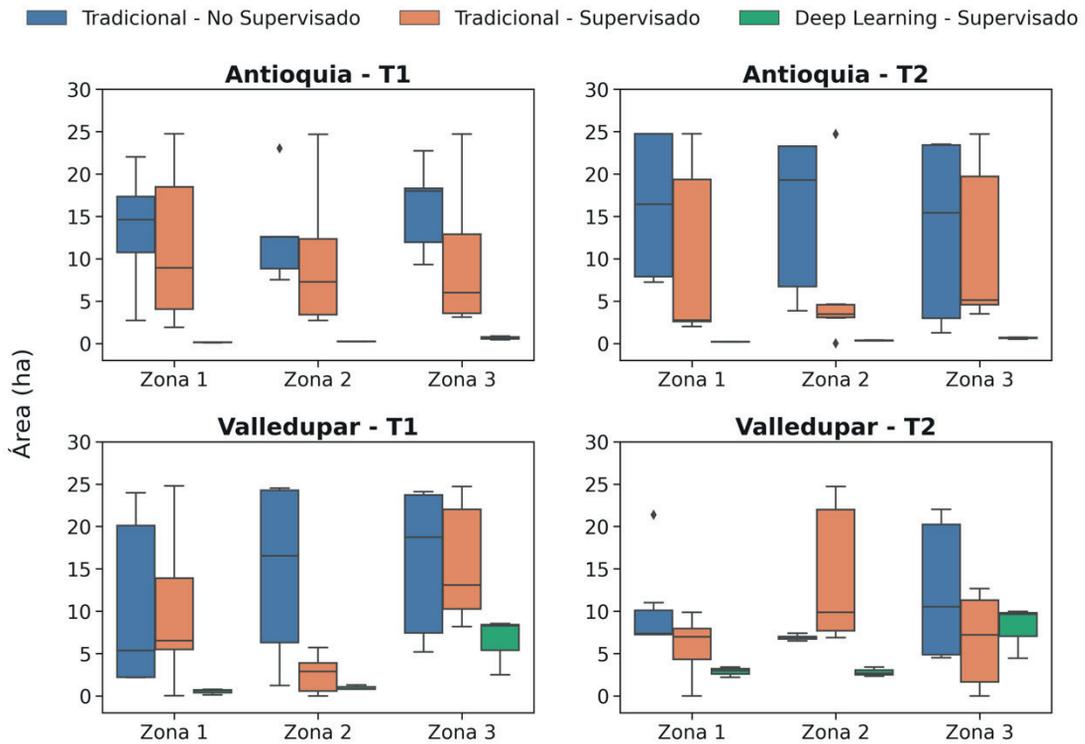


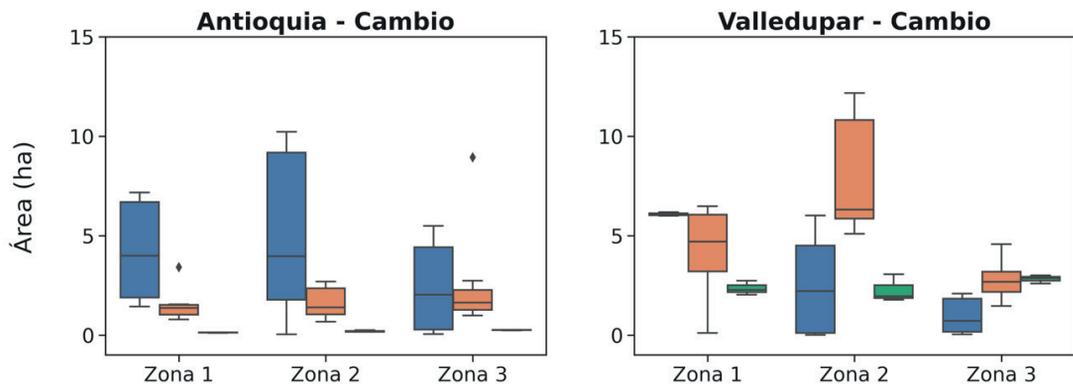
Figura 5. Gráficos de cajas comparando la funcionalidad según la métrica IoU de los algoritmos usados agrupados en tres familias en tres zonas de evaluación y dos tiempos de análisis en Antioquia y Valledupar.



**Figura 6.** Gráficos de caja comparando los valores de área de construcciones por parte de los algoritmos estudiados agrupados en tres familias para tres zonas de evaluación y dos tiempos de análisis en Antioquia y Valledupar.

Tabla 6 presenta los tres mejores ordenados de mayor a menor valor de IoU para las imágenes según su fecha, inicial (T1) y final (T2), así como el valor promedio. Adicionalmente, se muestran los valores de métricas derivadas de la matriz de confusión para cada fecha. Se observa que los algoritmos FPN con configuraciones de

entrenamiento combinado e individual ofrecen un aumento de IoU de 293% y 38% respecto al mejor de los tradicionales en los casos de Antioquia y Valledupar, respectivamente. A pesar de las ganancias reportadas por la red FPN, sus valores de IoU promedio de 0,40 y 0,54 en los ámbitos rural y urbano, respectivamente, se consideran



**Figura 7.** Gráficos de caja comparando los valores de área de cambio derivada de la diferencia de clasificación de los algoritmos estudiados agrupados en tres familias para tres zonas de evaluación y dos tiempos de análisis en Antioquia y Valledupar.

relativamente inferiores a los reportados por otros estudios con datos de entrada y ambientes similares. Por ejemplo, Gram-Hansen et al. (2019) entrenaron el modelo DeepLabv3+, un tipo de red convolucional, a partir de muestras con ruido (con una revisión mínima de los errores de topología) y bandas RGB de imágenes de alta resolución (0,3-0,5 m) para identificar construcciones ilegales en ocho ubicaciones distribuidas en Asia (India), África (Kenia, Nigeria y Sudán) y América del Sur (Colombia). Los valores de IoU variaron entre 0,65 y 0,83, siendo los menores valores obtenidos en ambientes rurales de Al Geneina, Sudán y los altos en ámbitos urbanos en Medellín, Colombia. Wang y Li (2020) propusieron una red basada en U-Net con conexiones residuales y módulos multiescala Atrous Spatial Pyramid Pooling (ASSP) para identificar construcciones en ámbitos urbanos de acuerdo a cuatro conjuntos de datos representando imágenes de satélite y drones con resoluciones entre 0,1 m y 4 m. De acuerdo a las muestras de evaluación, la red propuesta alcanzó valores de IoU entre 0,79 y 0,91, los cuales resultaron superiores a los obtenidos en otras redes evaluadas, entre ellas DeepLabv3+ (valores de IoU entre 0,70 y 0,75). Aunque los resultados de la red FPN en términos de la métrica IoU se consideran bajos, en especial en los ámbitos urbanos (donde con mayor proporción se tiene este tipo de estudios), los resultados pueden considerarse aceptables en términos de F1-Score para los ámbitos rurales. El estudio de Vargas-Muñoz et al. (2019), realizado a partir de muestras de OSM, reporta valores de F1-score entre 0,11 y 0,66, y 0,52 y 0,82 para identificar construcciones rurales en Tanzania y Zimbabue, respectivamente. Mientras que los valores bajos corresponden a las predicciones directas de la red, los valores altos se obtuvieron con el post-procesamiento de estas predicciones usando los campos aleatorios de Markov. El rango de valores, en especial para Tanzania, se aproxima a los obtenidos usando la red FPN en Antioquia (F1-Score de 0,69) (Tabla 6).

Es importante mencionar que la brecha observada entre la red FPN y el mejor método tradicional en los ámbitos rurales como Antioquia se podría reducir mediante el ingreso de una mayor cantidad de información espectral. Es común para los algoritmos supervisados tradicionales usar índices específicos que resaltan construcciones como el índice de áreas construidas de diferencia

normalizada (NDBI por sus siglas en inglés) (Zha et al., 2003). No obstante, estos requieren la banda del infrarrojo de onda corta, siendo esta fundamental para evitar las confusiones de suelo desnudo y superficies de área construida comunes en los ámbitos rurales (Lyu et al., 2018). Adicionalmente, a pesar de que en el entorno rural se puede usar una máscara de vías, como se realizó en esta investigación, aún no resulta suficiente para separar las construcciones de otros usos. Caso contrario, en un entorno urbano, la aplicación de dicho filtro resulta más favorable ya que está compuesto principalmente por vías, edificaciones y, en menor medida, parques, haciendo que la detección de las construcciones sea buena. La limitación de la adición de la banda del infrarrojo no existe por parte de los modelos entrenados FPN ya que fueron capaces de delimitar construcciones usando las bandas del azul, verde y rojo de las imágenes de alta resolución. Esto resalta el potencial de los modelos FPN para su implementación a gran escala, en este caso nacional, utilizando imágenes de drones y/o aerofotografías que usualmente carecen de las bandas del infrarrojo.

Una inspección visual de los mejores algoritmos por familia permite un mayor entendimiento del alcance de cada uno. Las Tablas 7 y 8 proveen las imágenes, vectores de referencia y resultados de mapa de clasificación y cambio en dos de las tres zonas de evaluación en Antioquia y Valledupar, respectivamente. Este ejercicio confirma que el modelo FPN ofrece el mejor resultado, con un menor ruido que los modelos tradicionales. Cabe indicar que el entrenamiento de esta red con el modelo preentrenado de EfficientNetB3 resultó favorable como fue evidenciado por Parmar et al. (2020). De acuerdo con una comparación preliminar (no presentada en este artículo), esta estrategia fue la que mejores resultados generó para la tarea de identificación y detección de cambios en construcciones. A diferencia de entrenar la arquitectura FPN desde su inicio con pesos aleatorios, con el uso del modelo preentrenado, se obtuvieron resultados razonables con un menor tiempo de entrenamiento.

Aunque los resultados obtenidos usando la plataforma GEE son promisorios, bajo el enfoque de comparación post-clasificación existen errores que se propagan en cada una de las escenas

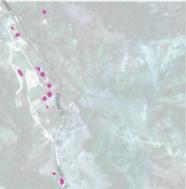
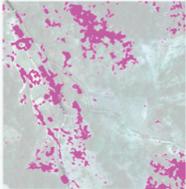
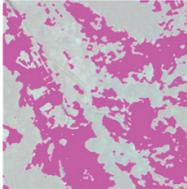
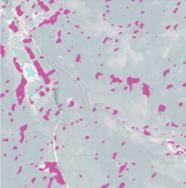
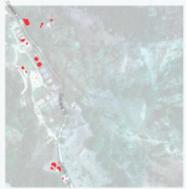
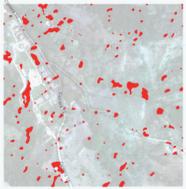
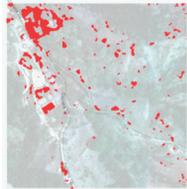
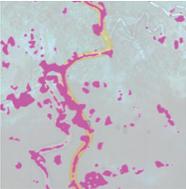
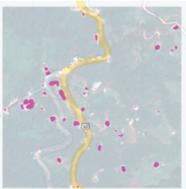
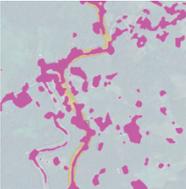
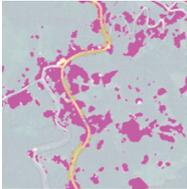
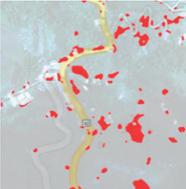
**Tabla 6.** Comparación de los mejores algoritmos por familia por área de estudio de acuerdo al IoU y cuatro métricas derivadas de la matriz de confusión: *Overall Accuracy (OA)*, *recall*, *precision* y *F1-score* por cada tiempo de clasificación. La columna de porcentaje de diferencia es calculada de acuerdo al mayor IoU promedio de ambas fechas, en este caso obtenido con los algoritmos *deep learning* en ambas áreas.

Área de estudio	Familia	Algoritmo	Métrica	Tiempo 1	Tiempo 2	IoU promedio (% diferencia)
Antioquia	<i>Deep Learning</i>	FPN efficientnetb3 - Vexcel/ Perusat-1 (Combinado)	IoU	0,39	0,41	0,40 (0)
			OA	0,99	0,98	
			Recall	0,70	0,69	
			Precision	0,68	0,70	
			F1-Score	0,69	0,69	
	Tradicional Supervisado	Support Vector Machine - Lineal	IoU	0,10	0,11	0,10 (293)
			OA	0,90	0,91	
			Recall	0,89	0,90	
			Precision	0,55	0,56	
			F1-Score	0,56	0,57	
	Tradicional No Supervisado	KMeans Manhattan	IoU	0,03	0,03	0,03 (1312)
			OA	0,76	0,73	
			Recall	0,66	0,77	
			Precision	0,51	0,51	
			F1-Score	0,45	0,45	
	Valledupar	<i>Deep Learning</i>	FPN efficientnetb3 – WorldView-2/Vexcel (Individual)	IoU	0,57	0,50
OA				0,92	0,86	
Recall				0,82	0,78	
Precision				0,83	0,77	
F1-Score				0,83	0,78	
Tradicional Supervisado		Random Forest	IoU	0,31	0,46	0,39 (38)
			OA	0,80	0,78	
			Recall	0,78	0,75	
			Precision	0,67	0,68	
			F1-Score	0,69	0,70	
Tradicional No Supervisado		KMeans Euclidean	IoU	0,37	0,37	0,37 (45)
			OA	0,79	0,74	
			Recall	0,66	0,63	
			Precision	0,62	0,61	
			F1-Score	0,63	0,61	

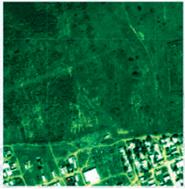
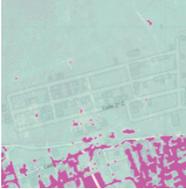
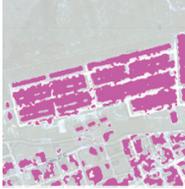
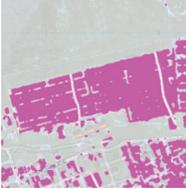
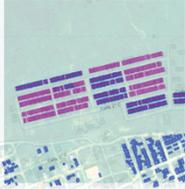
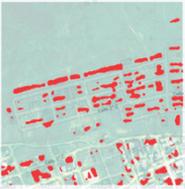
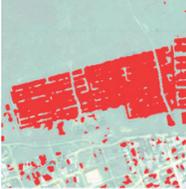
clasificadas, siendo este mayor cuando se manejan sensores diferentes por fecha. En este sentido, como trabajo futuro se plantea la exploración de técnicas de cambio directo que puedan manejar diferentes escalas y desplazamiento de imágenes de diversos sensores, ver por ejemplo Jiang et al. (2020). Igualmente, aunque en esta investigación se hizo un estudio sobre la correspondencia de los vectores, este proceso no es libre de errores. En este aspecto es común que las investigaciones planteen el uso de vectores de referencia corregidos y/o seleccionados sobre áreas libres de

errores topológicos y/o desplazamiento (Li et al., 2020). En ámbitos rurales, la baja exactitud en los vectores de entrada puede ser más severa que en las zonas urbanas, ya que las construcciones son más pequeñas y dispersas. En este aspecto, como se describió en la introducción, sería interesante evaluar la estrategia propuesta por Vargas-Muñoz et al. (2019), la cual usa las salidas de la red neuronal convolucional y post-procesamiento con campos aleatorios de Markov para corregir de manera automatizada etiquetas de construcciones de OSM en ámbitos rurales. Lo anterior puede

**Tabla 7.** Resultados visuales de los mejores algoritmos por familia en ambos tiempos de clasificación (píxeles magenta) y detección de cambio (píxeles rojos) en la zona de evaluación 1 y 2 de Antioquia. Los valores de área de construcciones con sus respectivos valores de IoU para cada tiempo de análisis así como el área de cambio son reportados en la parte inferior de cada figura.

Zona	Imagen RGB y vectores referencia	FPN efficientnetb3 - Vexcel/Perusat-1 (Combinado)	Support Vector Machine Lineal	KMeans Manhattan
1				
	Tiempo 1	0,12 ha / IoU 0,35	2,86 ha / IoU 0,03	10,30 ha / IoU 0,01
				
	Tiempo 2	0,21 ha / IoU 0,31	2,03 ha / IoU 0,08	7,23 ha / IoU 0,02
				
	Cambio - Sin cambio	0,12 ha	1,52 ha	1,43 ha
2				
	Tiempo 1	0,25 ha / IoU 0,48	2,72 ha / IoU 0,12	7,53 ha / IoU 0,03
				
	Tiempo 2	0,35 ha / IoU 0,55	2,46 ha / IoU 0,10	3,87 ha / IoU 0,09
				
	Cambio - Sin cambio	0,16 ha	3,87 ha	1,79 ha

**Tabla 8.** Resultados visuales de los mejores algoritmos por familia en ambos tiempos de clasificación (píxeles magenta) y detección de cambio (píxeles rojos) en la zona de evaluación 1 y 3 de Valledupar. Los valores de área de construcciones con sus respectivo valores de IoU para cada tiempo de análisis así como el área de cambio son reportados en la parte inferior de cada figura.

Zona	Imagen RGB y vectores referencia	FPN efficientnetb3 World-View2/Vexcel (Individual)	Random Forest	Kmeans Euclidean
1				
	Tiempo 1	0,76 ha / IoU 0,61	9,85 ha / IoU 0,08	2,21 ha / IoU 0,25
				
	Tiempo 2	2,99 ha / IoU 0,48	6,40 ha / IoU 0,41	7,34 ha / IoU 0,31
				
	Cambio - Sin cambio		2,27 ha	7,43 ha
3				
	Tiempo 1	8,53 ha / IoU 0,64	9,45 ha / IoU 0,67	5,34 ha / IoU 0,76
				
	Tiempo 2	9,65 ha / IoU 0,74	8,41 ha / IoU 0,73	6,66 ha / IoU 0,41
				
	Cambio - Sin cambio		2,60 ha	2,41 ha

considerarse como trabajo futuro complementario a las etiquetas oficiales como fue implementado en esta investigación.

## 5. Conclusiones

La presente investigación evaluó diferentes familias de algoritmos de clasificación de imágenes implementados en la plataforma GEE con el objetivo de identificar construcciones y su cambio en dos tiempos de análisis, considerando dos áreas de estudio representativas de ámbitos rurales y límites del perímetro urbano municipal en Colombia. Los resultados indican el potencial de la arquitectura FPN de la *familia deep learning* supervisado para esta tarea. Gracias a la reciente integración de GEE y Google AI, esta investigación es pionera en el uso de estas tecnologías para evaluar modelos de segmentación semántica como FPN usando imágenes de alta resolución para los casos colombianos. Igualmente, se contribuyó al estado del arte de estos modelos con enfoque en ámbitos rurales, los cuales no cuentan con una alta documentación comparada con las investigaciones en ámbitos urbanos. Finalmente, cabe indicar que el desarrollo de la aplicación en GEE permitió la experimentación en las zonas de evaluación, siendo esto una ventaja para poder replicar el estudio en otras áreas del país o regiones de condiciones similares donde estén disponibles los insumos de entrada.

## Agradecimientos

Los autores agradecen a las Subdirecciones de Catastro, y Geografía y Cartografía del IGAC. Esta investigación hace parte de la licencia del programa GEO-GEE administrada por la Subdirección de Geografía y Cartografía. Se agradece igualmente al equipo de EODataScience por su soporte constante en los desarrollos técnicos de esta investigación.

## Referencias

Abraham, N., Khan, N.M. 2019. A Novel Focal Tversky Loss Function With Improved Attention U-Net for Lesion Segmentation. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. Venecia, Italia, 8-11 Abril. pp. 683–687. <https://doi.org/10.1109/ISBI.2019.8759329>

- Chaurasia, A., Culurciello, E. 2018. LinkNet: Exploiting encoder representations for efficient semantic segmentation. *2017 IEEE Visual Communications and Image Processing, VCIP 2017*. St. Petersburg, USA, 10-13 Diciembre. <https://doi.org/10.1109/VCIP.2017.8305148>
- Daudt, R.C., Le Saux, B., Boulch, A., Gousseau, Y. 2018. Urban change detection for multispectral earth observation using convolutional neural networks. *International Geoscience and Remote Sensing Symposium (IGARSS)*. Valencia, España, 22-27 Julio. pp. 2115–2118. <https://doi.org/10.1109/IGARSS.2018.8518015>
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>
- Gram-Hansen, B.J., Helber, P., Varatharajan, I., Azam, F., Coca-Castro, A., Kopackova, V., Bilinski, P. 2019. Mapping informal settlements in developing countries using machine learning and low resolution multi-spectral data. *AIES 2019 - Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. Honolulu, USA, 27-28 Enero. <https://doi.org/10.1145/3306618.3314253>
- Han, Y., Javed, A., Jung, S., Liu, S. 2020. Object-Based Change Detection of Very High Resolution Images by Fusing Pixel-Based Change Detection Results Using Weighted Dempster–Shafer Theory. *Remote Sensing*, 12(6), 983. <https://doi.org/10.3390/rs12060983>
- He, H., Zhou, J., Chen, M., Chen, T., Li, D., Cheng, P. 2019. Building Extraction from UAV Images Jointly Using 6D-SLIC and Multiscale Siamese Convolutional Networks. *Remote Sensing*, 11(9), 1040. <https://doi.org/10.3390/rs11091040>
- Hussain, M., Chen, D., Cheng, A., Wei, H., Stanley, D. 2013. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80, 91–106. <https://doi.org/10.1016/j.isprsjprs.2013.03.006>
- Instituto Geográfico Agustín Codazzi - IGAC. 2020. Resolución 388, <https://www.igac.gov.co/es/noticias/resolucion-388-del-13-de-abril-de-2020>
- Jaccard, P. 1908. Nouvelles recherches sur la distribution florale. *Bulletin de La Société Vaudoise Des Sciences Naturelles*, 44(163), 223-270.
- Jiang, H., Hu, X., Li, K., Zhang, J., Gong, J., Zhang, M. 2020. PGA-SiamNet: Pyramid feature-based attention-guided siamese network for remote sensing orthoimagery building change detection. *Remote Sensing*, 12(3), 1–21. <https://doi.org/10.3390/rs12030484>

- Kingma, D.P., Ba, J. (2014, December 22). Adam: A Method for Stochastic Optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. San Diego, USA, 7-9 Mayo.
- Li, L.-J., Li, K., Li, F.F., Deng, J., Dong, W., Socher, R., Fei-Fei, L. 2009. ImageNet: a Large-Scale Hierarchical Image Database Shrimp Project View project hybrid intrusion detection systems View project ImageNet: A Large-Scale Hierarchical Image Database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*. Miami, USA, 20-25 Junio.
- Li, Q., Shi, Y., Huang, X., Zhu, X.X. 2020. Building Footprint Generation by Integrating Convolution Neural Network With Feature Pairwise Conditional Random Field (FPCRF). *IEEE Transactions on Geoscience and Remote Sensing*, 1–18. <https://doi.org/10.1109/TGRS.2020.2973720>
- Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S. 2017. Feature pyramid networks for object detection. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. Honolulu, USA, 21-26 Julio. pp 2117-2125. <https://doi.org/10.1109/CVPR.2017.106>
- Liu, Y., Pang, C., Zhan, Z., Zhang, X., Yang, X. 2019. *Building Change Detection for Remote Sensing Images Using a Dual Task Constrained Deep Siamese Convolutional Network Model*. Recuperado de: arXiv. Último acceso: 22 de Diciembre, 2020, de <http://arxiv.org/abs/1909.07726>
- Lyu, H., Lu, H., Mou, L., Li, W., Wright, J., Li, X., Li, X., Zhu, X., Wang, J., Yu, L., Gong, P. 2018. Long-Term Annual Mapping of Four Cities on Different Continents by Applying a Deep Information Learning Method to Landsat Data. *Remote Sensing*, 10(3), 471. <https://doi.org/10.3390/rs10030471>
- Ma, L., Li, M., Blaschke, T., Ma, X., Tiede, D., Cheng, L., Chen, Z., Chen, D. 2016. Object-based change detection in Urban Areas: The effects of segmentation strategy, scale, and feature space on unsupervised methods. *Remote Sensing*, 8(9), 1–18. <https://doi.org/10.3390/rs8090761>
- Parmar, V., Bhatia, N., Negi, S., Suri, M. 2020. Exploration of Optimized Semantic Segmentation Architectures for edge-Deployment on Drones. Recuperado de: *arXiv*. Último acceso: 18 de Abril, 2021, <http://arxiv.org/abs/2007.02839>
- Pinzón-Rodríguez, O.A., Sanabria-García, J.C. 2017. *Determinación de zonas geoeconómicas para el proyecto "Construcción, rehabilitación, operación y mantenimiento y revisión del sistema vial para la conexión de los departamentos del Cesar y la Guajira"*. Universidad Distrital Francisco José de Caldas. <http://hdl.handle.net/11349/7752>
- Ronneberger, O., Fischer, P., Brox, T. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *MICCAI: Medical Image Computing and Computer-Assisted Intervention*. Munich, Alemania, 5-9 Octubre. pp. 234–241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- Rußwurm, M., Körner, M. 2018. Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders. *ISPRS International Journal of Geo-Information*, 7(4), 129. <https://doi.org/10.3390/ijgi7040129>
- Vargas-Muñoz, J.E., Lobry, S., Falcão, A.X., Tuia, D. 2019. Correcting rural building annotations in OpenStreetMap using convolutional neural networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 147, 283-293. <https://doi.org/10.1016/j.isprsjprs.2018.11.010>
- Vásquez, J., Smith, R., Cadena, Á., Rave, C. 2006. Planificación energética local bajo una evaluación integrada energía-ambiente-economía. Caso de aplicación valle de aburra y canasta energética en distribución de energía de Empresa Públicas de Medellín. *Energética*, 36, 25–33.
- Wang, C., Li, L. 2020. Multi-scale residual deep network for semantic segmentation of buildings with regularizer of shape representation. *Remote Sensing*, 12(18), 2932. <https://doi.org/10.3390/rs12182932>
- Yang, H.L., Yuan, J., Lunga, D., Laverdiere, M., Rose, A., Bhaduri, B. 2018. Building Extraction at Scale Using Convolutional Neural Network: Mapping of the United States. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), 2600–2614. <https://doi.org/10.1109/JSTARS.2018.2835377>
- Zha, Y., Gao, J., Ni, S. 2003. Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing*, 24(3), 583–594. <https://doi.org/10.1080/01431160304987>
- Zhang, L., Wu, J., Fan, Y., Gao, H., Shao, Y. 2020. An efficient building extraction method from high spatial resolution remote sensing images based on improved mask R-CNN. *Sensors (Switzerland)*, 20(5), 1–13. <https://doi.org/10.3390/s20051465>
- Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J. 2017. Pyramid scene parsing network. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*. Honolulu, USA, 21-26 Julio. pp 6230-6239. <https://doi.org/10.1109/CVPR.2017.660>