

Change Detection over Mining Areas using Deep Neural Networks from PlanetScope Image Time Series

M.A. Zaraza-Aguilera, E. Upegui, *Member, IEEE* and O.J. Espejo-Valero

Abstract— The processes of change detection (CD) in land use and land cover through satellite images acquired in different temporal phases represent a key process to monitor the land, the environment and evaluate disasters. Currently the methodologies used for CD based on pixels do not allow processing a large amount of images massively, so it is not possible to analyze all the available information. Main aim of this research is to evaluate the performance of Deep learning methods for CD, grouped in two approaches, across open mining areas from PlanetScope (PS) image time series of the Cerrejón mine in Colombia. Two approaches for change generation are proposed, one based on post-classification comparison in which two convolutional neural network (CNN) architectures are evaluated: U-Net and Feature Pyramid Network (FPN) for the classification of mining areas along the time series, for this purpose different models with different hyperparameters were generated and trained to select the most suitable for such process and subsequently perform the difference between the periods of the series; and a second approach based on direct change detection in which a modified U-Net network was evaluated from pairs of images. For this purpose, different models were also trained selecting the most appropriate for the detection of changes for each period of the time series, obtaining a map of changes for each approach, each one the results were validated; the most appropriate approach was number 2 (Direct change detection), with kappa accuracies greater than 0.9 in each period of the time series.

Index Terms—Change Detection, Deep Learning, Mining areas, PlanetScope, Time Series

I. INTRODUCCION

La detección de cambios (DC) es uno de los principales temas de investigación en teledetección, y es esencial para el procesamiento y la comprensión de los datos de observación de la Tierra disponibles a gran escala [1]. La DC es el proceso de identificar las diferencias en el estado de un objeto o fenómeno observado en diferentes momentos y proporciona información para el monitoreo urbano, la gestión de recursos, la agricultura, la evaluación y la planificación de los ecosistemas y el medio ambiente, así como el monitoreo oportuno de los desastres naturales, entre otros [2].

Los métodos tradicionales para DC se basan principalmente en la diferencia de imágenes, análisis de componentes principales y comparaciones post-clasificación [2]. En el proceso de detección se comparan y analizan los cambios entre diferentes imágenes satelitales y datos relacionados en diferentes periodos de tiempo a través de algoritmos y modelos matemáticos. El principal objetivo de utilizar imágenes satelitales para detectar cambios en la superficie terrestre, es dar respuesta a tres preguntas: ¿Existen cambios?, ¿Cuál es el cambio?, ¿Cuándo se dio el cambio?, para ello se han empleado métodos de álgebra de imágenes, clasificaciones, transformación, modelos híbridos, uso de sistemas de información geográfica, análisis visual, etc.[3].

Con el aumento del volumen de datos o imágenes satelitales obtenidas por distintos programas satelitales, algunos métodos tradicionales o convencionales se van limitando, principalmente por dos aspectos; el primero relacionado con la variedad de datos, ya que el mejoramiento de la resolución espacial en las imágenes y el enriquecimiento de los detalles en la imagen conllevan a que métodos como el álgebra de imágenes sean fácilmente influenciados por los efectos de cambios estacionales, sensor del satélite, elevación solar y condiciones atmosféricas que reducen la exactitud en la DC [4]. El segundo aspecto está relacionado con los métodos orientados a objetos y transformaciones, que generan falsos cambios por la extracción geométrica y la textura, además requiere de largos tiempos de procesamiento y procesos manuales, lo que limita las ventajas de las técnicas de procesamiento automático [3]. Es por ello que nace la demanda y la necesidad por una mayor capacidad técnica en el análisis de grandes volúmenes de datos multispectrales que permitan generar mapas temáticos cada vez más exactos y oportunos sobre cambios en las coberturas, con una menor pérdida de información y la mayor cobertura tanto espacial como temporal [5] y con tiempos de procesamiento más cortos a los obtenidos con los métodos tradicionales.

En los últimos años, las arquitecturas de aprendizaje profundo (Deep Learning - DL) se han convertido en un método novedoso y popular en teledetección para tareas de clasificación y detección de cambios [6], aportando altas precisiones y procesando la información espacio-temporal de forma más eficiente frente a los métodos tradicionales [7], [8]. Las redes neuronales convolucionales (CNNs) son los modelos o arquitecturas más ampliamente implementados en DC [9]. Varios estudios, han podido determinar la mejora en la exactitud de la identificación y clasificación de coberturas y áreas de cambio frente a los métodos tradicionales, la utilización y aprovechamiento de una gran cantidad de datos de diferentes sensores y con un análisis más riguroso por su

M.A. Zaraza-Aguilera, Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, mazarazaa@correo.udistrital.edu.co.

E. Upegui, Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, esupeguic@udistrital.edu.co.

O.J. Espejo-Valero, Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, ojespejov@correo.udistrital.edu.co.

aprendizaje profundo y detallado de la información no solo a nivel espectral sino teniendo en cuenta la forma y distribución de los elementos [7]–[15].

Las series temporales de imágenes de satélite (STIS) proporcionan una poderosa herramienta para aprender sobre eventos pasados, monitorear las condiciones actuales y pronosticar cambios futuros [16]. Pocos estudios han analizado el proceso de detección de cambios mediante STIS en áreas de minería, mientras que el uso de datos bitemporales permite detectar cambios a partir de un par de imágenes, el análisis de STIS ofrece un contexto más detallado de los cambios a lo largo de diferentes temporalidades.

Esta investigación tiene como objetivo principal evaluar el rendimiento de los métodos de Deep Learning para DC sobre áreas de minería a cielo abierta en STIS de imágenes Planet Scope (PS) en la mina del Cerrejón para el periodo 2017 -2019, agrupados en dos esquemas. El primer esquema se basa en la comparación post-clasificación, en el que se evaluaron dos métodos: i) redes de segmentación basadas en CNN bajo las arquitecturas U-Net y FPN; el segundo enfoque corresponde a una detección directa de cambios para lo cual se evaluó una red U-Net modificada y entrenada directamente a partir de muestras de cambios entre pares de imágenes. Una vez obtenidos los resultados de los diferentes modelos bajo los dos esquemas se busca escoger el más adecuado para la DC sobre áreas de minería.

II. TRABAJOS RELACIONADOS

En esta sección se analizan los trabajos de investigación relacionados detrás de los métodos propuestos para la detección de cambios en la cobertura vegetal.

A. Deep Learning

En los últimos años, la utilización de algoritmos de DL en diferentes aplicaciones ha aumentado de manera importante, encontrando resultados favorables en diversos campos como medicina [17], [18]; geociencias, [19], [20]; detección de objetos [21], entre otros.

En 2019, se realizó una revisión y análisis de más de 200 publicaciones recientes en el campo de DL, en relación al tema de detección de cambios en las coberturas de la tierra, encontrando que las series de tiempo son un área de aplicación prometedora para clasificación y detección de cambios en la cobertura vegetal [7].

Algunas de las aplicaciones se han enfocado en la detección y clasificación de la minería, como el trabajo realizado por Gallwey y su equipo [22] para la detección de minería artesanal en Ghana a partir de imágenes satelitales multispectrales Sentinel-2 alcanzando una precisión superior al 80%; o la investigación realizada por Balaniuk, Isupova & Reece para la búsqueda de minas de oro y presas no registradas en grandes áreas de Brasil [23] a partir de imágenes Sentinel-2, donde demuestran que estas pueden ser identificadas y clasificadas automáticamente con una precisión del 95%.

B. Detección de Cambios

La aplicación de los métodos de DC permitan generar mapas temáticos cada vez más oportunos sobre cambios en las

coberturas, con una menor pérdida de información y mayor cobertura espacio-temporal, y una mejor resolución espacial. Entre las técnicas empleadas para la DC se encuentran las que usan imágenes continuas, por ejemplo, aritmética entre imágenes, vector de cambios, regresión, o análisis de componentes principales, entre otros; y las que usan imágenes discretas en las que comparan imágenes previamente clasificadas a través de la clasificación directa, o comparación post-clasificación [24]. Algunos trabajos también incluyen el uso de metodologías híbridas. Los principales procesos de detección de cambios implementando DL se basan en diferencia de imágenes (DI) [25]. Una de las grandes ventajas del uso de redes neuronales profundas es la capacidad de emplear y analizar datos de múltiples fuentes (Ópticas, Radar, DEM, etc.) y de diferentes periodos de tiempos combinando varias técnicas de fusión, lo cual ayuda a mejorar la exactitud de los procesos de detección de cambios [26].

C. U-Net

U-Net es una de las redes neuronales profundas de segmentación semántica más conocidas. Fue inicialmente creada para analizar imágenes biomédicas, sin embargo, tiene cada vez más relevancia en casi todas las áreas de conocimiento, incluyendo la teledetección, en esta arquitectura, el codificador se utiliza para la extracción de características en múltiples niveles y el decodificador combina las características y resolución aprendidas a través de un sofisticado apilamiento, teniendo en cuenta la localización y representación de las características [23]. La investigación de Balaniuk, Isupova & Reece [23] utiliza esta red para la detección de deforestación en la amazonia de Brasil mediante imágenes Landsat 8; y el trabajo de Gallwey y su equipo [22] la ha empleado para detección y clasificación de áreas de minería artesanal a cielo abierto en Ghana a partir de imágenes Sentinel-2.

D. Feature Pyramid Networks (FPN)

La red FPN trabaja creando dos pirámides, y las combina para generar mapas de segmentación resaltando las características en cada nivel [23], este método toma como entrada una imagen a escala única de tamaño arbitrario y produce mapas de características de tamaño proporcional en múltiples niveles, de forma totalmente convolucional [23]. La arquitectura se compone de capas multiescala producidas a partir de una estructura piramidal de abajo-arriba (bottom-up), arriba-abajo (top-down) y por el uso de conexiones laterales, adicionalmente, debido su enfoque de capas multi-escala, las FPN permiten la adición de redes pre-entrenadas en la estructura abajo-arriba. Esto último favorece los tiempos de entrenamiento reduciendo considerablemente el uso de recursos.

Las anteriores arquitecturas fueron empleadas porque han sido ampliamente utilizadas en diversos estudios entre ellos el realizado por Parmar y su equipo [27] para la clasificación de imágenes de dron, en este estudio las redes U-Net y FPN junto con el codificador EfficientNetB3 obtuvieron los mejores resultados. Adicional a este estudio, también se destaca el realizado por Castro y otros [28] en el que evaluaron dichas dos

arquitecturas con resultados bastante satisfactorios para la segmentación de áreas urbanas; por lo que para este proyecto se decidió evaluar la eficacia de la red FPN y U-Net con inicialización de pesos usando el modelo preentrenado de EfficientNetB3, para la detección de cambios por minería.

III. AREA DE ESTUDIO Y DATOS

El área de estudio corresponde a una región de la Mina a cielo abierto del Cerrejón ubicada en el Departamento de la Guajira (Colombia), en la que se extrae carbón y materiales de construcción, el área de estudio abarca un área de 327.16 km² comprendiendo 5 municipios pertenecientes al departamento (Riohacha, Barrancas, Hato Nuevo, Albania y Maicao). La mina está ubicada en la cuenca del río Ranchería, al sureste del Departamento, al este de la Sierra Nevada de Santa Marta y al oeste de la Serranía del Perijá, en la línea de frontera con Venezuela, Fig. 1.

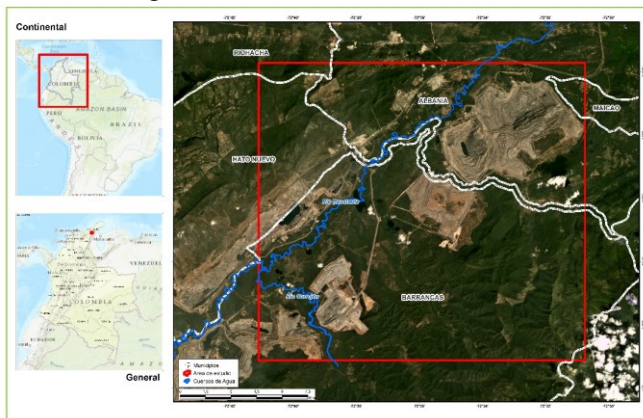


Fig. 1. Área de estudio Mina el Cerrejón.

A. Conjunto de Datos

PS opera con más de 175 satélites que recolectan imágenes multiespectrales en 4 bandas (Azul (B), Verde (G), Rojo (R) e Infrarrojo Cercano-IRC) con una capacidad de fotografiar toda el área terrestre cada día [29]. Para la serie de tiempo se emplearon 8 imágenes distribuidas entre octubre de 2017 y noviembre de 2019 seleccionadas con el menor porcentaje de nubosidad sobre la zona de estudio, estas imágenes poseen un nivel de procesamiento 3B; las cuales son orto rectificadas, con una alta exactitud y se encuentran en valores de reflectancia de superficie, cuentan con una resolución espacial de 3 metros. El entrenamiento y configuración de los modelos evaluados se realizó bajo la versión 2.1 de Tensorflow y en el entorno en la nube de Google Cloud Colab, con acceso libre a una Unidad Gráfica de Procesamiento (GPU) Tesla K80.

IV. METODOLOGIA

La metodología planteada se dividió en cinco fases (Fig. 2); la primera asociada a la preparación de los datos raster y vector, se recortaron los datos a la extensión del área de estudio con un conjunto de 8 imágenes PS y 3 capas vectoriales que corresponden a áreas de minería y no minería de diciembre de 2017, 2018 y 2019; la segunda fase está asociada a la generación de muestras las cuales tienen en cuenta el contexto

del píxel o parche, para extraer las áreas de entrenamiento, validación y evaluación, para esto se generó una grilla de bloques de 1.6 km x 1.6 km con una separación de 0.4 km (aplicando la estrategia usada en investigaciones anteriores [30], [31]).

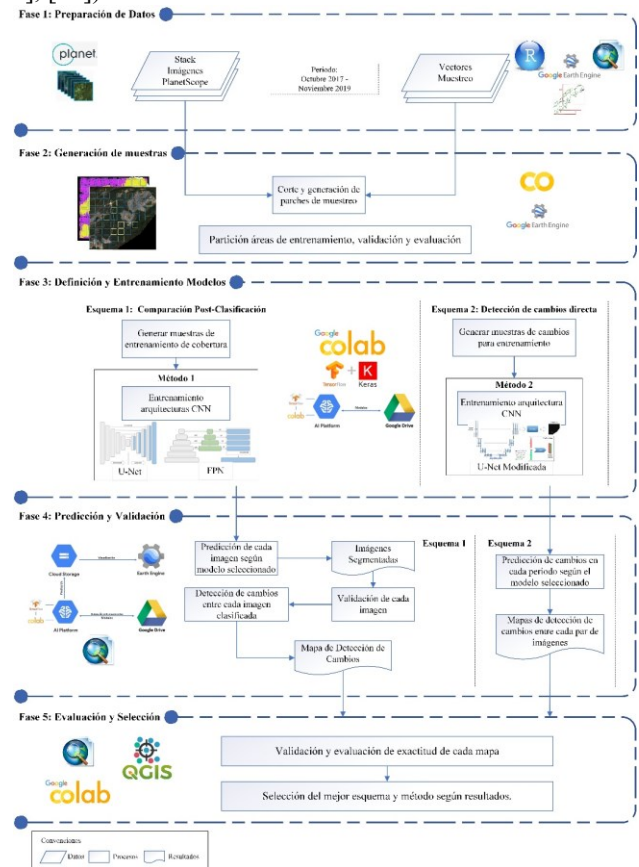


Fig. 2. Metodología para la detección de áreas de cambio por minería.

La Fase 3 corresponde a la definición y entrenamiento de los modelos bajo dos esquemas. El primer esquema está basado en la comparación post-clasificación (Fig. 3A) el cual consiste en la generación de clasificaciones de cada una de las imágenes que conforman la STIS y posteriormente se realiza una diferencia entre los vectores de clasificación entre par de fechas de la serie de tiempo, se evaluaron dos tipos de arquitecturas CNN: (i) U-Net, [32] y (ii) FPN, [33].

Se entrenaron un total de 12 modelos con 3 canales (RGB), 4 canales (RGBIRC) y 5 canales (RGBIRC mas el Índice de Vegetación de Diferencia Normalizada (NDVI) [34]), con diferente número de épocas (50 y 100); las funciones de pérdida empleadas fueron Dicefocal, [35] y Focal-Tversky, [36] para U-Net y FPN, respectivamente. Este tipo de métrica representa la suma de errores del modelo y permite medir que tan bien el modelo se está entrenando [37]. De los 12 modelos se escogen los más adecuados para ser reentrenados con datos de las tres temporalidades (diciembre 2017, 2018 y 2019) y con variación en el número de épocas.

El segundo esquema corresponde a detección directa de cambios (Fig. 3B), empleando una red U-Net modificada [38], la red se entrena a partir de observaciones de cambio entre pares de imágenes, incorpora tanto el contexto espacial como el

temporal. Este último permite reducir los falsos positivos, así como aumentar el enfoque en los cambios; las áreas de cambio se obtienen directamente entre cada par de imágenes de la serie temporal. Un total de 5 modelos fueron entrenados bajo este esquema, con variaciones en el número de épocas y número de bandas (4 y 5 canales) y la función de pérdida Dicefocal.

Para cada modelo se empleó un tamaño de lote (batch size) de 16, como indica el trabajo de Ball, Anderson & Chan [39] este es un valor adecuado ya que valores más grandes tienden a tiempos de entrenamiento más largos; el optimizador Adam, [40] el cuál es el más adoptado en los modelos de segmentación; la tasa de aprendizaje definida fue de 0,1 como en otros estudios [27], [28].

Para evaluar el rendimiento de los modelos en ambos esquemas en las etapas de entrenamiento, validación y evaluación, se utilizó el índice intersección sobre unión (IoU – Intersection over Union) [41], el cual mide el coeficiente de similitud de la clase objetivo entre dos áreas, valores cercanos a 1 indican una alta correspondencia entre la predicción y el área de referencia. También se empleó la medida Loss que indica la desviación entre las predicciones realizadas por el modelo y los valores reales a partir de la función de pérdida, [42], valores cercanos a 0 indican que el modelo es eficiente.

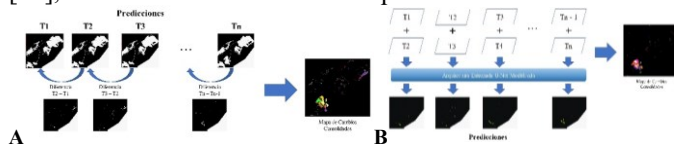


Fig. 3. Esquemas A. Comparación post-clasificación B. Detección directa.

La Fase 4 corresponde a la predicción y generación de los cambios detectados en cada uno de los periodos de la serie de tiempo para cada uno de los dos esquemas. La Fase 5 corresponde a la validación y evaluación de exactitud de capa mapa de cambios la cual se realizó mediante la metodología planteada por Olofsson y su equipo [43], la cual aborda principalmente tres fases: (i) un diseño de muestreo probabilístico; (ii) un diseño de la respuesta que sea coherente con la información de referencia a usar; y (iii) un análisis para estimar la exactitud, a partir una matriz de error; finalmente se selecciona el esquema más adecuado para el proceso de detección de cambios de minería en la serie de tiempo.

V. RESULTADOS Y DISCUSION

En esta sección se presentan los resultados obtenidos en cada fase propuesta, según los dos esquemas planteados para la obtención de las áreas de cambio bajo los algoritmos de DL.

A. Resultados Fase 2: Generación de Muestras

En total, 57 bloques cubrieron el área de estudio, correspondientes a 42 para entrenamiento, 11 de validación y 4 para evaluación (Fig. 4). A partir del diseño por bloques, se generaron aleatoriamente parches de 256 píxeles x 256 píxeles sobre el conjunto de bloques, se utilizaron estrategias de aumento de datos como la rotación y el volteo horizontal y vertical. En total para el esquema 1 se generaron 11050, 2700 y 1050 parches para entrenamiento, validación y evaluación de los modelos respectivamente. Para el esquema 2 se obtuvieron

8150, 1950 y 800 parches para entrenamiento, validación y evaluación respectivamente.

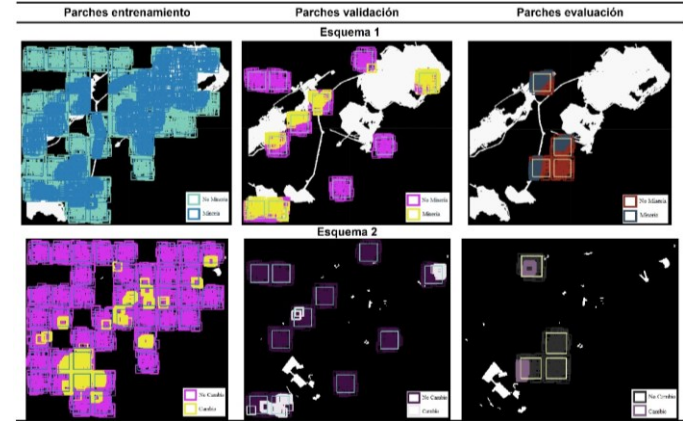


Fig. 4. Bloques y parches de entrenamiento, validación y evaluación.

B. Resultados Fase 3: Definición y Entrenamiento de Modelos.

A nivel general los 12 modelos bajo el esquema 1 en el entrenamiento poseen métricas de exactitud (IoU) superiores a 0.98 (TABLA I) mientras que con los datos de validación los modelos disminuyen su exactitud levemente rodeando valores alrededor de 0.9. Con los datos de evaluación los valores en general son buenos los cuales se encuentran entre 0.75 y 0.81 a excepción del modelo 6 el cual obtuvo un IoU de 0.72 esto puede indicar que un modelo con dicha configuración requiere de más de 50 épocas para tener una mayor exactitud ya que el modelo 5 el cual posee las mismas características, pero con 100 épocas posee un IoU de 0.79. Se evidencia una relación inversa entre los valores de Loss y IoU en cada conjunto de datos, ya que con valores altos de Loss, el IoU disminuye tanto en la arquitectura FPN como U-Net. Entre los modelos con la red FPN se destacan el número 4 (4 bandas - 50 épocas) y 5 (5 bandas - 100 épocas) mientras que con la red U-Net los mejores modelos corresponden al 10 y 11 con las mismas características de los de la red FPN cuyos valores para las áreas de evaluación se encuentran entre 0.79 y 0.81, estos cuatro modelos fueron los escogidos para ser reentrenados (TABLA II - Fig. 5).

TABLA I
MEDIDAS LOSS Y IOU PARA LOS MODELOS DEL ESQUEMA 1.

Id	Modelo	Entrenamiento		Validación		Evaluación	
		Loss	IoU	Loss	IoU	Loss	IoU
1	FPN_3b_100e	0.12	0.99	0.29	0.91	0.46	0.75
2	FPN_3b_50e	0.14	0.99	0.30	0.90	0.45	0.76
3	FPN_4b_100e	0.12	0.99	0.32	0.88	0.48	0.75
4	FPN_4b_50e	0.14	0.99	0.30	0.89	0.41	0.81
5	FPN_5b_100e	0.11	0.99	0.31	0.88	0.42	0.79
6	FPN_5b_50e	0.14	0.98	0.30	0.90	0.49	0.72
7	UNET_3b_100e	0.01	0.99	0.13	0.90	0.21	0.76
8	UNET_3b_50e	0.01	0.98	0.14	0.88	0.35	0.75
9	UNET_4b_100e	0.01	0.99	0.15	0.89	0.22	0.75
10	UNET_4b_50e	0.02	0.97	0.12	0.89	0.16	0.79
11	UNET_5b_100e	0.01	0.99	0.17	0.89	0.17	0.81
12	UNET_5b_50e	0.02	0.98	0.14	0.88	0.20	0.76

Una vez reentrenados los modelos, se destaca el 4 con un IoU de 0.74 (TABLA II), sin embargo a nivel visual (Fig. 5) se

evidencia que los modelos poseen un alto grado de acierto ya que logran hacer una identificación acertada de la zona de minería y una delimitación detallada de la misma. En algunos modelos como el 2 y 4 (5 bandas) poseen algunas áreas clasificadas como minería pero que en realidad no lo son, mientras que en el número 3 las áreas mal clasificadas son mínimas a nivel visual, logrando una consistencia en la forma y delimitación del área minera y su diferenciación con coberturas aledañas como suelos, zona urbana o vegetación. Con base en los resultados, se puede indicar que el modelo más adecuado para la predicción de las áreas mineras bajo el esquema 1 sobre la serie de tiempo es el número 3 correspondiente a la arquitectura U-Net con la función de pérdida DiceFocal, con 4 bandas (RGB-NIR) y con 75 épocas.

Entre las dos arquitecturas FPN y U-Net las diferencias son mínimas a nivel de los resultados en cada modelo, hay una mayor consistencia con la forma y delimitación de las zonas asociadas a minería bajo la red U-Net lo que lo llevo a ser escogido como el más adecuado. Sin embargo, la red FPN reentrenada con mayor cantidad de datos puede refinar su comportamiento y predicción. A nivel estructural la diferencia entre ambas redes se da en que la red FPN aplica una capa de convolucion 1x1 antes de añadir los datos a la siguiente capa mientras que U-Net solo copia las características y las añade [44], esto permite que U-Net genere una menor deformación de los datos de entrenamiento y la red tenga un mayor entendimiento de la información ingresada.

La elección de la red U-Net como la mejor frente a la red FPN contrasta con lo propuesto por Parmar y otros [27], donde la arquitectura FPN comparada con otras como U-Net, Linknet y PSPNet fue la más óptima para la segmentación semántica de coberturas sobre una imagen tomada por UAV. Allí el análisis detallado de la arquitectura FPN aporta un mejor resultado al tener en cuenta la variabilidad de las coberturas. Otros estudios como el de Castro y otros [28] también demuestran el potencial de la arquitectura FPN para la clasificación de construcciones en imágenes satelitales Worldview-2 frente a otros clasificadores de machine learning como máquinas de soporte vectorial. En este caso se demuestran las ventajas de la red U-Net para la clasificación de áreas de minería frente a la red FPN.

TABLA II
MEDIDAS LOSS Y IOU PARA LOS MEJORES MODELOS DEL ESQUEMA 1.

Id	Modelo	Entrenamiento		Validación		Evaluación	
		Loss	IoU	Loss	IoU	Loss	IoU
1	FPN_4b_50e	0.15	0.98	0.31	0.89	0.49	0.71
2	FPN_5b_75e	0.13	0.99	0.31	0.89	0.50	0.70
3	UNET_4b_75e	0.01	0.99	0.15	0.89	0.33	0.70
4	UNET_5b_75e	0.01	0.98	0.17	0.89	0.29	0.74

Los buenos resultados de la arquitectura U-Net se ven respaldados por otras investigaciones como la de Gallwey y otros [22] en la que la emplean para la segmentación de áreas mineras de pequeña escala en Ghana en imágenes Sentinel-2 obteniendo un IoU en validación de 0.86 y entrenada con 30 épocas y la función de pérdida entropía cruzada binaria, ya que encontraron que la función Dicefocal tiende a un entrenamiento inestable y de baja convergencia. Sin embargo, en comparación a lo obtenido en este trabajo se evidencia que esta función de pérdida permite dar un ajuste adecuado al modelo evitando el

sobreajuste del mismo, a su vez el IoU obtenido es un poco más alto (0.88) por lo tanto el modelo aquí estructurado permite aportar a una posible configuración de este tipo de arquitecturas para la segmentación de áreas de minería.

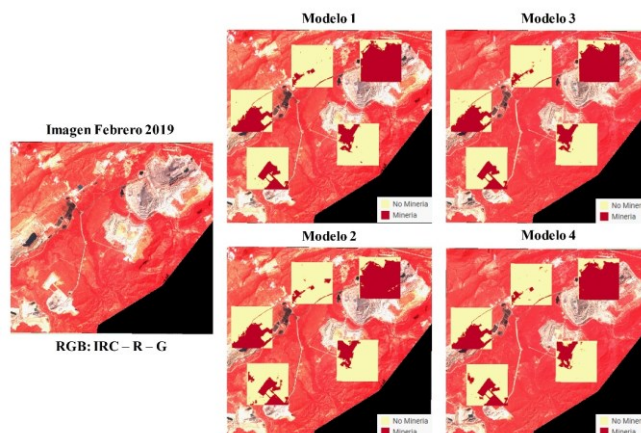


Fig. 5. Predicciones de los 4 modelos reentrenados bajo el esquema 1.

Bajo el esquema 2, los 5 modelos entrenados (TABLA III) obtuvieron exactitudes un poco más bajas a las obtenidas bajo el esquema 1 en las áreas de evaluación; el modelo entrenado bajo un tamaño de lote de 32, el cual corresponde al número de ejemplos utilizados para estimar el gradiente de error antes de actualizar los parámetros del modelo y a su vez controlar la rapidez con la que un modelo aprende [42]. Bajo este modelo se evidenció que un valor alto indica mayor velocidad de aprendizaje, pero a su vez más generalización en el entendimiento de los datos, no hay una optimización y estabilidad del modelo y a su vez el modelo no está realizando un ajuste acertado con respecto al comportamiento de los datos. Por lo tanto, los valores de pérdida (loss) en validación y evaluación son más altos frente a los demás modelos.

Los modelos 1 y 2 poseen valores de IoU más altos a nivel de evaluación frente a los modelos de 5 bandas (3 y 4). En la Fig. 6 se evidencia la predicción de cambios detectados entre la imagen de diciembre de 2017 y noviembre de 2019, donde el modelo 2 es el más consistente, acertado con la delimitación y el que menos errores de omisión o comisión presenta a nivel visual. En los demás modelos los polígonos tienen algunos pixeles clasificados como no cambio dentro de las áreas de cambio por lo que baja la consistencia de dichos polígonos y a su vez genera algunas áreas mal detectadas.

TABLA III
MEDIDAS LOSS Y IOU PARA LOS MEJORES MODELOS DEL ESQUEMA 2.

Id	Modelo	Entrenamiento		Validación		Evaluación	
		Loss	IoU	Loss	IoU	Loss	IoU
1	UNET_4b_40e	0.14	0.88	0.60	0.48	0.21	0.74
2	UNET_4b_45e	0.16	0.88	0.59	0.46	0.22	0.73
3	UNET_5b_20e	0.18	0.78	0.73	0.33	0.35	0.61
4	UNET_5b_25e	0.17	0.82	0.65	0.40	0.36	0.59
5	UNET_5b_50e_b32	0.24	0.74	1.01	0.63	0.45	0.59

En el esquema 2 se puede indicar que el modelo más adecuado para la predicción de las áreas de cambio es el número 2 correspondiente a la red U-Net modificada con la función de pérdida DiceFocal, con 4 bandas (RGB-NIR) y con 45 épocas.

Este modelo obtuvo áreas concisas visualmente con una delimitación bastante detallada en cuanto a la forma y bordes de las zonas de cambio y una buena separabilidad con respecto a las áreas de no cambio.

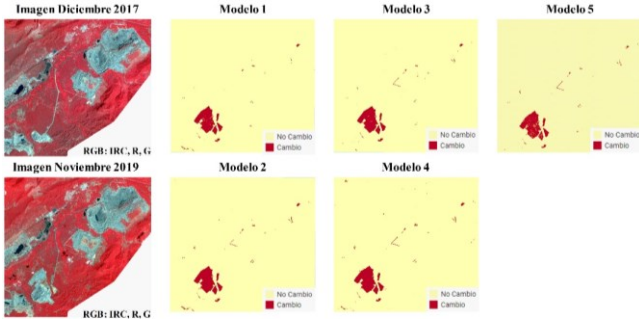


Fig. 6. Predicciones bajo los 5 modelos entrenados en el esquema 2.

C. Resultados Fase 4: Predicción y Validación

A partir de los modelos escogidos se generan las segmentaciones para cada imagen de la serie de tiempo bajo el esquema 1 y las áreas de cambios en ambos esquemas.

En general las predicciones obtenidas bajo el primer esquema en cada una de las imágenes que conforman la serie de tiempo son acordes con lo evidenciado en la Fig. 7. Se logran identificar bastante bien el área correspondiente a minería, los diferentes frentes de explotación ubicados en la zona son delimitados correctamente al igual que las vías que los conectan, se excluyen aquellas zonas de vegetación que se encuentran dentro de los tajos de las áreas de extracción y se incluyen las piscinas empleadas en las labores de explotación minera. Existen leves errores en áreas aledañas a la zona urbana en zonas que fueron identificadas como minería, también es de destacar que en áreas de nube (Áreas A y B - Fig. 7), el modelo logra evitar la confusión y no las identifica como minería, tal como se evidencia en la imagen de octubre de 2017 donde en el frente de explotación de mayor tamaño (Área A) hay un área faltante de la zona minera dado por la presencia de nubes y sombras. Se puede indicar que el modelo empleado para la predicción a nivel visual posee resultados bastantes acertados con la imagen de cada fecha.

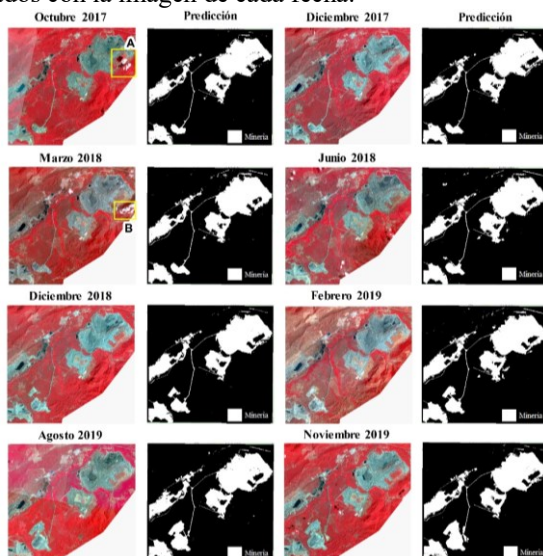


Fig. 7. Segmentaciones obtenidas bajo el esquema 1 a lo largo de la serie de tiempo. Combinaciones imágenes RGB: (IRC, R, G)

A partir de las imágenes predichas por cada fecha de la serie de tiempo y el esquema planteado en la Fig. 3A, se extraen los cambios mediante comparaciones entre cada par de imágenes con el fin de obtener las áreas nuevas correspondientes a minería, dadas por la ampliación de un frente de explotación o aparición de nuevas áreas asociadas a actividades mineras. En el proceso de análisis no se tuvieron en cuenta las áreas recuperadas o que pasaron de minería a vegetación o suelo.

En general las variaciones en las áreas de cambio son bajas entre los periodos (Fig. 8), las áreas de cambio se encuentran en un rango entre 140 y 228 has. Se identifican algunas áreas que no corresponden a cambios por lo tanto hay errores de comisión en la estimación de los mismos que llevan a una leve sobredimensión en las áreas calculadas, los principales cambios se dan en el frente de explotación ubicado en la parte inferior izquierda (Área 1 - Fig. 8) el cual va ampliando su área a lo largo de los diferentes periodos. Los otros cambios identificados con mayor área se presentan en la zona oriental (Área 2) dado por las variaciones en la clasificación donde algunas imágenes fueron afectadas por nubosidad, en otras la clasificación delimita áreas propias de los tajos y en otras las generalizo, por lo tanto, al hacer el proceso de diferencia se identificaron cambios no correspondientes a minería.

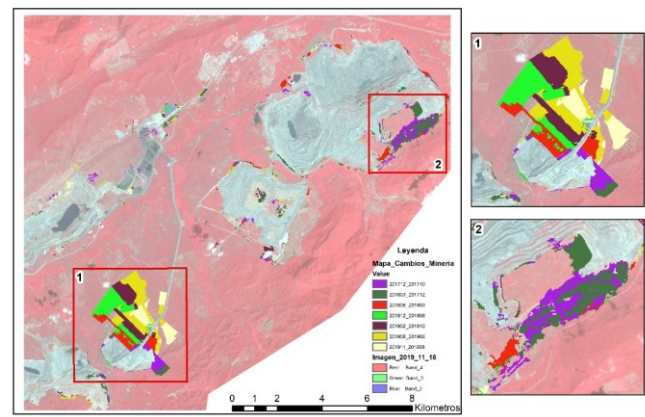


Fig. 8. Mapa de cambios obtenidas bajo el esquema 1.

Bajo el esquema 2 se obtienen los cambios directamente de cada par de imágenes que conforman la serie de tiempo, los cambios detectados al igual que en el esquema anterior corresponden a la ampliación de un frente de explotación o la creación o aparición de nuevas áreas (Área 1 - Fig. 9). Con respecto a las áreas de cambio obtenidas en este esquema, se evidencia un descenso en los cambios a lo largo de los 3 periodos iniciales (Fig. 10) con respecto al esquema 1, las mayores ampliaciones se dan en el frente de explotación del área 1. El penúltimo periodo es el de mayor área de cambios (250 has) con ampliaciones del frente de explotación indicado nuevas áreas de minería alrededor de vías y de nuevos tramos de las mismas que conectan algunos sectores de los frentes de explotación (Recuadro 2 - Fig. 9).

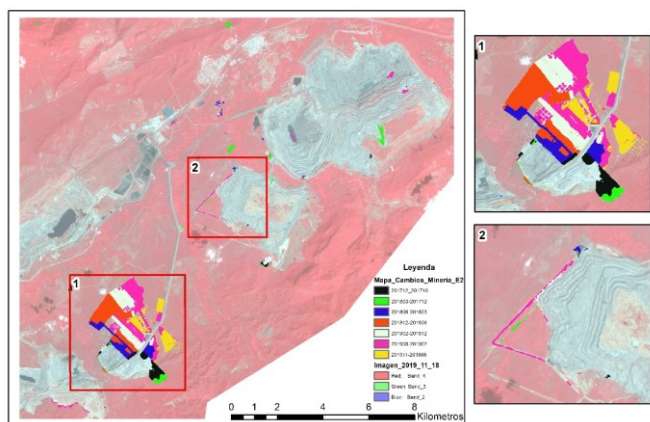


Fig. 9. Mapa de cambios obtenidos bajo el esquema 2.

D. Resultados Fase 5: Evaluación y Selección.

En la TABLA IV se encuentran las medidas de exactitud obtenidas para las áreas de Cambio (Cmb) y No Cambio (No Cmb) en cada periodo (Pr) de la serie de tiempo bajo el esquema 1. A nivel de la exactitud de usuario (Ex. Usuario) de cambios se tienen valores inferiores a 0.8 en los primeros tres periodos siendo el más bajo el primero con 0.62, lo que a su vez indica que los errores de comisión se encuentran alrededor de 0.3 esto a lleva a que exista una sobredimensión en las áreas de cambio en dichos periodos. En los periodos posteriores donde las imágenes no están afectadas por nubosidad la exactitud mejora considerablemente lo que indica que los cambios tienen un alto porcentaje de correspondencia con la realidad. Con respecto a la exactitud del productor (Ex. Productor) las métricas poseen valores altos lo que indica que los errores por omisión son bajos las exactitudes se encuentran entre 0.87 y 0.97 siendo la más baja en el primer periodo.

A nivel del coeficiente kappa la menor exactitud se da en los dos primeros periodos con valores alrededor de 0.7, en los periodos posteriores este va aumentando considerablemente indicando la alta concordancia entre los cambios detectados con la realidad.

TABLA IV

MEDIDAS DE EVALUACIÓN PARA LOS CAMBIOS OBTENIDOS EN EL ESQUEMA 1.

Pr	Rango	Ex. Usuario		Ex. Productor		Kappa
		No Cmb	Cmb	No Cmb	Cmb	
1	201712_201710	0.99	0.62	0.87	0.99	0.70
2	201803_201712	0.99	0.68	0.89	0.97	0.74
3	201806_201803	0.99	0.79	0.93	0.99	0.84
4	201812_201806	0.99	0.89	0.96	0.98	0.92
5	201902_201812	0.99	0.83	0.94	0.98	0.87
6	201908_201902	0.99	0.93	0.97	0.99	0.95
7	201911_201908	0.99	0.83	0.94	0.99	0.87

En el esquema 2 la exactitud de usuario obtenida demarca valores altos en cada periodo (TABLA V) superiores a 0.93 para la clase de cambio a excepción del segundo periodo que tiene un valor de 0.8 lo cual indica un 0.2 de error de comisión. Para el caso de la exactitud del productor se obtuvieron valores altos en ambas clases superiores a 0.9 lo cual indica que los

errores de omisión son bajos. Con respecto al coeficiente kappa los valores son superiores a 0.92 en todos los periodos exceptuando el segundo con un coeficiente kappa de 0.84. Los altos valores en general de la validación reflejan los buenos resultados del clasificador en áreas de cambio y evidenciando leves errores de omisión y comisión; los resultados poseen una alta correspondencia con la realidad según lo evidenciado en cada par de imágenes que conforman la serie de tiempo.

TABLA V

MEDIDAS DE EVALUACIÓN PARA LOS CAMBIOS OBTENIDOS EN EL ESQUEMA 2.

Pr	Ex. Usuario		Ex. Productor		Kappa
	No Cmb	Cmb	No Cmb	Cmb	
1	0.98	0.93	0.97	0.97	0.93
2	0.99	0.80	0.93	0.96	0.84
3	0.99	0.95	0.98	0.98	0.96
4	0.99	0.97	0.98	0.97	0.96
5	0.98	0.98	0.99	0.96	0.97
6	0.99	0.97	0.99	0.98	0.97
7	0.99	0.93	0.97	0.98	0.94

Los resultados bajo el esquema 2 son bastante buenos comparados con los obtenidos en otros como el planteado por Maretto y su equipo [38] en el cual emplean dos tipos de U-Net modificadas, la primera corresponde a una fusión de las dos temporalidades antes de iniciar la codificación de la red, la segunda U-Net (tardía) codifica cada imagen y después de la convolucion son combinadas. Las redes fueron entrenadas con 100 épocas y la función de pérdida entropía cruzada por pesos y una variación de la Dicefocal para la identificación de áreas deforestadas en la amazonia Brasileña bajo imágenes Landsat-8 [38]; de lo cual obtienen una exactitud f1-score de 0.95 en cada red, levemente mayor al obtenido en este trabajo (kappa promedio 0.94 - TABLA V), lo cual puede estar dado por la resolución de las imágenes al emplear un mayor detalle y minimizar la generalización.

A partir de los resultados de cambios obtenidos bajo cada esquema para la zona de minería (TABLA IV-TABLA V), se tiene que el de mayor exactitud en términos del coeficiente kappa es el esquema 2 ya que los valores en su mayoría son superiores a 0.9 (TABLA V) y existe una mejor identificación de los elementos con mínimos errores de omisión y comisión. En este esquema no hubo afectación de las áreas de nubosidad presentes en el esquema 1, las cuales generaron confusión con algunas áreas de minería dificultando la separabilidad de las mismas y generando un sobredimensionamiento de las áreas de cambios, como indica la Fig. 10 donde los primeros tres periodos en el esquema 1 obtuvieron mayor área que en el esquema 2. En los demás periodos las áreas de cambio son similares entre ambos esquemas con mínimas variaciones, dichas variaciones en las áreas de cambio se ven reflejadas en los coeficientes kappa de ambos esquemas. Con base en lo anterior se puede indicar que el esquema más apropiado para la detección, clasificación y cuantificación de cambios de coberturas por minería a lo largo de la serie de tiempo analizada es el número 2.

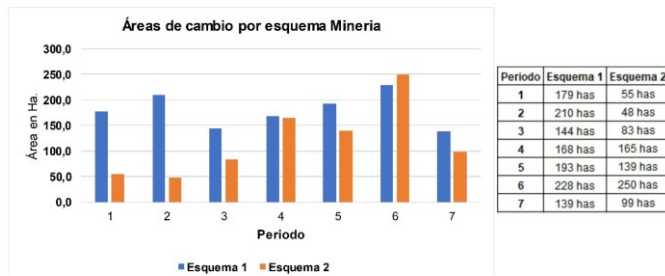


Fig. 10. Áreas de cambio obtenidas por cada esquema.

VI. CONCLUSIONES.

Los esquemas planteados y evaluados (1-Comparación post-clasificación y 2-Detección de cambios directa), basados en algoritmos de DL permiten apoyar el proceso de DC en áreas de minería con mejores resultados a nivel temático, mejorando la exactitud en los procesos de detección, delimitación y clasificación de cambios de dichas coberturas de manera más oportuna y eficiente frente a los métodos tradicionales, con una buena delimitación de los elementos manteniendo su forma compacta y con bajos niveles de ruido.

El esquema número 2 fue el más apropiado para la detección y clasificación de cambios a lo largo de la serie de tiempo bajo una red U-Net modificada con la función de pérdida DiceFocal, con 4 bandas y 45 épocas obteniendo un IoU promedio a nivel de entrenamiento de 0.88 y evaluación 0.73. En los cambios generados se obtuvo un coeficiente kappa superior a 0.9 en cada periodo; sin embargo, cabe destacar que ambos esquemas aquí planteados y evaluados obtuvieron resultados bastante aceptables con diferencias mínimas entre ellos a nivel visual y de métricas.

Es de resaltar que los modelos entrenados con 4 bandas obtuvieron mejores resultados que los de 5, es decir al incluir el índice NDVI al entrenamiento de los mismos no hay un mayor aporte al mejoramiento y aumento de la exactitud en la detección e identificación de los cambios bajo los dos esquemas, por lo que no es una variable que aporte al proceso.

Bajo las dos arquitecturas CNN evaluadas en el esquema 1 (FPN y U-Net), se puede concluir que la arquitectura con mejores resultados a nivel de exactitud y de las zonas de evaluación predichas es la red U-Net, corroborando lo evaluado por Parmar y otros [27], y Castro y su equipo [28] donde la red FPN es más apropiada para coberturas o imágenes de mayor resolución mientras que la red U-Net es más apropiada en imágenes de media resolución ya que sus resultados fueron más consistentes y acordes con lo evidenciado a lo largo de la STIS.

Las métricas empleadas permitieron evaluar el grado de exactitud de los modelos y su grado de pérdida con los diferentes conjuntos de datos (entrenamiento, validación y evaluación) e identificar el sobreentrenamiento existente en algunos modelos. A su vez el coeficiente kappa permitió indicar cuál fue el esquema más adecuado en la DC de cada periodo.

Por último se destaca que la estructuración de los dos esquemas y de las arquitecturas evaluadas permitió construir un conjunto de datos referentes a minería bastante robustos para el entrenamiento, validación y evaluación de modelos de DL, los cuales pueden ser consultados en el siguiente repositorio:

<https://shortest.link/2wXm> junto con los principales códigos generados y empleados en las diferentes etapas del proceso, como insumo o guía a futuros proyectos que se puedan derivar del trabajo aquí propuesto.

REFERENCES

- [1] R. C. Daudt, B. L. Saux, A. Boulch, and Y. Gousseau, "Multitask Learning for Large-scale Semantic Change Detection," *ArXiv181008452 Cs*, Aug. 2019, Accessed: Jan. 25, 2021. [Online]. Available: <http://arxiv.org/abs/1810.08452>
- [2] X. Yuan, J. Tian, and P. Reinartz, "Building Change Detection Based on Deep Learning and Belief Function," in *2019 Joint Urban Remote Sensing Event (JURSE)*, May 2019, pp. 1–4. doi: 10.1109/JURSE.2019.8808968.
- [3] Q. Wang, X. Zhang, G. Chen, F. Dai, Y. Gong, and K. Zhu, "Change detection based on Faster R-CNN for high-resolution remote sensing images," *Remote Sens. Lett.*, vol. 9, no. 10, pp. 923–932, Oct. 2018, doi: 10.1080/2150704X.2018.1492172.
- [4] F. Pacifici, F. D. Frate, C. Solimini, and W. J. Emery, "An Innovative Neural-Net Method to Detect Temporal Changes in High-Resolution Optical Satellite Imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2940–2952, Sep. 2007, doi: 10.1109/TGRS.2007.902824.
- [5] M. C. Hansen *et al.*, "High-Resolution Global Maps of 21st-Century Forest Cover Change," *Science*, vol. 342, no. 6160, pp. 850–853, Nov. 2013, doi: 10.1126/science.1244693.
- [6] Y. Li, H. Zhang, X. Xue, Y. Jiang, and Q. Shen, "Deep learning for remote sensing image classification: A survey," *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, p. e1264, May 2018, doi: 10.1002/widm.1264.
- [7] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogramm. Remote Sens.*, vol. 152, pp. 166–177, Jun. 2019, doi: 10.1016/j.isprsjprs.2019.04.015.
- [8] M. Yang, L. Jiao, F. Liu, B. Hou, and S. Yang, "Transferred Deep Learning-Based Change Detection in Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, pp. 1–14, 2019, doi: 10.1109/TGRS.2019.2909781.
- [9] R. C. Daudt, B. L. Saux, A. Boulch, and Y. Gousseau, "Urban Change Detection for Multispectral Earth Observation Using Convolutional Neural Networks," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2018, pp. 2115–2118. doi: 10.1109/IGARSS.2018.8518015.
- [10] A. M. E. Amin, Q. Liu, and Y. Wang, "Convolutional neural network features based change detection in satellite images," in *First International Workshop on Pattern Recognition*, Jul. 2016, vol. 10011, p. 100110W. doi: 10.1117/12.2243798.
- [11] M. E. A. Arabi, M. S. Karoui, and K. Djerriri, "Optical Remote Sensing Change Detection Through Deep Siamese Network," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2018, pp. 5041–5044. doi: 10.1109/IGARSS.2018.8518178.
- [12] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, "Land Use Classification in Remote Sensing Images by Convolutional Neural Networks," *ArXiv150800092 Cs*, Aug. 2015, Accessed: Sep. 19, 2019. [Online]. Available: <http://arxiv.org/abs/1508.00092>
- [13] S. H. Khan, X. He, F. Porikli, and M. Bennamoun, "Forest Change Detection in Incomplete Satellite Images With Deep Neural Networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5407–5423, Sep. 2017, doi: 10.1109/TGRS.2017.2707528.
- [14] Y. Li, C. Peng, Y. Chen, L. Jiao, L. Zhou, and R. Shang, "A Deep Learning Method for Change Detection in Synthetic Aperture Radar Images," *IEEE Trans. Geosci. Remote Sens.*, pp. 1–13, 2019, doi: 10.1109/TGRS.2019.2901945.
- [15] G. J. Scott, M. R. England, W. A. Starns, R. A. Marcum, and C. H. Davis, "Training Deep Convolutional Neural Networks for Land-Cover Classification of High-Resolution Imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 549–553, Apr. 2017, doi: 10.1109/LGRS.2017.2657778.
- [16] W. J. D. van Leeuwen, B. J. Orr, S. E. Marsh, and S. M. Herrmann, "Multi-sensor NDVI data continuity: Uncertainties and implications for vegetation monitoring applications," *Remote Sens. Environ.*, vol. 100, no. 1, pp. 67–81, Jan. 2006, doi: 10.1016/j.rse.2005.10.002.

- [17] M. Tsunekii, "Deep learning models in medical image analysis," *J. Oral Biosci.*, Mar. 2022, doi: 10.1016/j.job.2022.03.003.
- [18] F. Behrad and M. Saniee Abadeh, "An overview of deep learning methods for multimodal medical data mining," *Expert Syst. Appl.*, vol. 200, p. 117006, Aug. 2022, doi: 10.1016/j.eswa.2022.117006.
- [19] W. Zhang, X. Gu, L. Tang, Y. Yin, D. Liu, and Y. Zhang, "Application of machine learning, deep learning and optimization algorithms in geoenvironment and geoscience: Comprehensive review and future challenge," *Gondwana Res.*, vol. 109, pp. 1–17, Sep. 2022, doi: 10.1016/j.gr.2022.03.015.
- [20] Z. Ma and G. Mei, "Deep learning for geological hazards analysis: Data, models, applications, and opportunities," *Earth-Sci. Rev.*, vol. 223, p. 103858, Dec. 2021, doi: 10.1016/j.earscirev.2021.103858.
- [21] K. Tong and Y. Wu, "Deep learning-based detection from the perspective of small or tiny objects: A survey," *Image Vis. Comput.*, vol. 123, p. 104471, Jul. 2022, doi: 10.1016/j.imavis.2022.104471.
- [22] J. Gallwey, C. Robiati, J. Coggan, D. Vogt, and M. Eyre, "A Sentinel-2 based multispectral convolutional neural network for detecting artisanal small-scale mining in Ghana: Applying deep learning to shallow mining," *Remote Sens. Environ.*, vol. 248, p. 111970, Oct. 2020, doi: 10.1016/j.rse.2020.111970.
- [23] R. Balaniuk, O. Isupova, and S. Reece, "Mining and Tailings Dam Detection in Satellite Imagery Using Deep Learning," *Sensors*, vol. 20, no. 23, Art. no. 23, Jan. 2020, doi: 10.3390/s20236936.
- [24] M. Hussain, D. Chen, A. Cheng, H. Wei, and D. Stanley, "Change detection from remotely sensed images: From pixel-based to object-based approaches," *ISPRS J. Photogramm. Remote Sens.*, vol. 80, pp. 91–106, Jun. 2013, doi: 10.1016/j.isprsjprs.2013.03.006.
- [25] G. Cao, B. Wang, H.-C. Xavier, D. Yang, and J. Southworth, "A new difference image creation method based on deep neural networks for change detection in remote-sensing images," *Int. J. Remote Sens.*, vol. 38, no. 23, pp. 7161–7175, Dec. 2017, doi: 10.1080/01431161.2017.1371861.
- [26] W. Shi, M. Zhang, R. Zhang, S. Chen, and Z. Zhan, "Change Detection Based on Artificial Intelligence: State-of-the-Art and Challenges," *Remote Sens.*, vol. 12, no. 10, Art. no. 10, Jan. 2020, doi: 10.3390/rs12101688.
- [27] V. Parmar, N. Bhatia, S. Negi, and M. Suri, "Exploration of Optimized Semantic Segmentation Architectures for edge-Deployment on Drones," *ArXiv200702839 Cs*, Jul. 2020, Accessed: Jan. 12, 2021. [Online]. Available: <http://arxiv.org/abs/2007.02839>
- [28] A. C. Castro *et al.*, "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," *Rev. Teledetec. Rev. Asoc. Esp. Teledetec.*, no. 58, pp. 71–88, 2021.
- [29] Planet, "PLANET IMAGERY PRODUCT SPECIFICATIONS." 2021. [Online]. Available: https://assets.planet.com/docs/Planet_Combined_Imagery_Product_Specs_letter_screen.pdf
- [30] A. Coca-Castro, M. Rußwurm, L. Reymondin, and M. Mulligan, "Sequential Recurrent Encoders for Land Cover Mapping in The Brazilian Amazon Using Modis Imagery and Auxiliary Datasets," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2019, pp. 9843–9846. doi: 10.1109/IGARSS.2019.8899114.
- [31] M. Rußwurm and M. Körner, "Multi-Temporal Land Cover Classification with Sequential Recurrent Encoders," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 4, p. 129, Mar. 2018, doi: 10.3390/ijgi7040129.
- [32] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *ArXiv150504597 Cs*, May 2015, Accessed: Jan. 25, 2021. [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [33] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature Pyramid Networks for Object Detection," *ArXiv161203144 Cs*, Apr. 2017, Accessed: Feb. 02, 2021. [Online]. Available: <http://arxiv.org/abs/1612.03144>
- [34] D. Vallauri, J. Aronson, N. Dudley, and R. Vallejo, "Monitoring and Evaluating Forest Restoration Success," in *Forest Restoration in Landscapes: Beyond Planting Trees*, S. Mansourian, D. Vallauri, and N. Dudley, Eds. New York, NY: Springer, 2005, pp. 150–158. doi: 10.1007/0-387-29112-1_21.
- [35] W. Zhu *et al.*, "AnatomyNet: Deep Learning for Fast and Fully Automated Whole-volume Segmentation of Head and Neck Anatomy," *Med. Phys.*, vol. 46, no. 2, pp. 576–589, Feb. 2019, doi: 10.1002/mp.13300.
- [36] N. Abraham and N. M. Khan, "A Novel Focal Tversky Loss Function With Improved Attention U-Net for Lesion Segmentation," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, Apr. 2019, pp. 683–687. doi: 10.1109/ISBI.2019.8759329.
- [37] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.
- [38] R. V. Maretto, L. M. G. Fonseca, N. Jacobs, T. S. Körting, H. N. Bendini, and L. L. Parente, "Spatio-Temporal Deep Learning Approach to Map Deforestation in Amazon Rainforest," *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2020, doi: 10.1109/LGRS.2020.2986407.
- [39] J. E. Ball, D. T. Anderson, and C. S. Chan, "Comprehensive survey of deep learning in remote sensing: theories, tools, and challenges for the community," *J. Appl. Remote Sens.*, vol. 11, no. 4, p. 042609, Sep. 2017, doi: 10.1117/1.JRS.11.042609.
- [40] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic optimization," 2015.
- [41] M. A. Rahman and Y. Wang, "Optimizing Intersection-Over-Union in Deep Neural Networks for Image Segmentation," in *Advances in Visual Computing*, Cham, 2016, pp. 234–244. doi: 10.1007/978-3-319-50835-1_22.
- [42] J. Brownlee, "A Gentle Introduction to the Challenge of Training Deep Learning Neural Network Models," *Machine Learning Mastery*, Feb. 14, 2019. <https://machinelearningmastery.com/a-gentle-introduction-to-the-challenge-of-training-deep-learning-neural-network-models/> (accessed Nov. 29, 2020).
- [43] P. Olofsson, G. M. Foody, M. Herold, S. V. Stehman, C. E. Woodcock, and M. A. Wulder, "Good practices for estimating area and assessing accuracy of land change," *Remote Sens. Environ.*, vol. 148, pp. 42–57, May 2014, doi: 10.1016/j.rse.2014.02.015.
- [44] T. S. Sharan, S. Tripathi, S. Sharma, and N. Sharma, "Encoder Modified U-Net and Feature Pyramid Network for Multi-class Segmentation of Cardiac Magnetic Resonance Images," *IETE Tech. Rev.*, vol. 0, no. 0, pp. 1–13, Aug. 2021, doi: 10.1080/02564602.2021.1955760.



Maycol Zaraza Aguilera received the degree Cadastral and Geodesy Engineer in 2017 his specialization in Geographic Information System in 2019 from Universidad Distrital Francisco Jose de Caldas, Bogota, Colombia. He is currently developing his M.Sc. studies in Information and Communications Sciences in the same institution. His research interests are focused in the field of satellite image processing, machine learning, CD and time series.



Erika Upegui (M'2015) Associate Professor for the School of Engineering at Universidad Distrital Francisco José de Caldas. She obtained her PhD title in Geography and Territorial Planning at Université Franche-Comté in France in 2012. She is MSc in Teledetection and Geomatics Applied to Environmental Sciences at Université Paris 7 (France -2009). She received the degree Cadastral and Geodesy Engineer in 2002 at Universidad Distrital Francisco José de Caldas (Colombia).



Javier Espejo Born in Bogotá, received the Cadastral and Geodesy Engineer B.S in 2004 and M.S in Information and Communications Sciences in 2012 from the Universidad Distrital Francisco Jose de Caldas. His research interests are focused in the field of GEOBIA, LULC, LULUCF and Crops Monitoring Systems.