

AmazonCRIME: un conjunto de datos y punto de referencia de Inteligencia Artificial Geoespacial para la clasificación de áreas potenciales vinculadas a Crímenes Ambientales Transnacionales en la Selva Amazónica

Jairo J. Pinto-Hidalgo *, Jorge A. Silva-Centeno

Programa de Posgraduación en Ciencias Geodésicas, Universidade Federal do Paraná, Av. Coronel Francisco Heráclito dos Santos, 210, Jardim das Américas, Curitiba, Brasil.

Resumen: En este artículo es abordado el desafío de detectar áreas vinculadas con crímenes ambientales transnacionales en la selva amazónica usando datos de Inteligencia Geoespacial, imágenes de libre acceso Sentinel-2 proporcionadas por el programa Copernicus, así como también las capacidades de procesamiento en la nube de la plataforma *Google Earth Engine*. Para esto, se generó un conjunto de datos que consta de 6 clases con un total de 30.000 imágenes multiespectrales de 13 bandas, etiquetadas y georreferenciadas que es usado para alimentar modelos avanzados de Inteligencia Artificial Geoespacial (redes neuronales convolucionales profundas) especializados en las tareas de clasificación de imágenes. Con el conjunto de datos presentado en este artículo es posible obtener una exactitud global (*overall accuracy*) de clasificación de 96,56%. Es también demostrado cómo los resultados obtenidos se pueden utilizar en aplicaciones reales para apoyar la toma de decisiones destinadas a prevenir los Crímenes Ambientales Transnacionales en la selva Amazónica. El Conjunto de datos *AmazonCRIME* se coloca a disposición del público en el repositorio: <https://github.com/jp-geoAI/AmazonCRIME.git>.

Palabras clave: Crímenes Ambientales Transnacionales, Selva Amazónica, Sentinel-2, Inteligencia Geoespacial, Inteligencia Artificial Geoespacial.

AmazonCRIME: a Geospatial Artificial Intelligence dataset and benchmark for the classification of potential areas linked to Transnational Environmental Crimes in the Amazon Rainforest

Abstract: In this article the challenge of detecting areas linked to transnational environmental crimes in the Amazon rainforest is addressed using Geospatial Intelligence data, open access Sentinel-2 imagery provided by the Copernicus programme, as well as the cloud processing capabilities of the Google Earth Engine platform. For this, a dataset consisting of 6 classes with a total of 30,000 labelled and geo-referenced 13-band multispectral images was generated, which is used to feed advanced Geospatial Artificial Intelligence models (deep convolutional neural networks) specialised in image classification tasks. With the dataset presented in this paper it is possible to obtain a classification overall accuracy of 96.56%. It is also demonstrated how the results obtained can be used in real

To cite this article: Pinto-Hidalgo, J.J., Silva-Centeno, J.A. 2022. AmazonCRIME: a Geospatial Artificial Intelligence dataset and benchmark for the classification of potential areas linked to Transnational Environmental Crimes in the Amazon Rainforest. *Revista de Teledetección*, 59, 1-21. <https://doi.org/10.4995/raet.2021.15710>

* Corresponding author: jairopinto@ufpr.br

applications to support decision making aimed at preventing Transnational Environmental Crimes in the Amazon rainforest. The AmazonCRIME Dataset is made publicly available in the repository: <https://github.com/jp-geoAI/AmazonCRIME.git>.

Key words: Transnational Environmental Crimes, Amazon rainforest, Sentinel-2, Geospatial Intelligence, Geospatial Artificial Intelligence.

1. Introducción

Los Crímenes Ambientales Transnacionales (CATs) se han convertido en uno de los principales impulsores financieros de la delincuencia organizada, afectan directamente los recursos naturales y generan serias consecuencias y amenazas para el mantenimiento de la paz, seguridad, salud humana, economía y desarrollo sostenible de la sociedad civil y los gobiernos (International Criminal Police Organization [INTERPOL], 2018). Se refieren a cualquier actividad u omisión en contra de la ley, cruzan las fronteras, se basan en el uso ilegal de los recursos naturales, originando impactos directos sobre el medio ambiente irreversibles y de grande alcance.

El carácter transnacional, los vincula con otras actividades delictivas, como el tráfico ilícito de drogas, legitimación de capitales, crimen organizado, corrupción y trata de personas, las cuales varían entre países y regiones geográficas (White, 2011). En general, realizan transformaciones en la diversidad biológica que contribuyen con el calentamiento global (INTERPOL, 2018; Zabyelina et al., 2020; Global Initiative Against Transnational Organized Crime, 2021). En algunos casos, están fuertemente estructurados por grupos del crimen organizado, grandes corporaciones, gobiernos cómplices, funcionarios públicos corruptos y profesionales altamente especializados (Gore et al., 2019), utilizan técnicas sofisticadas, tecnologías modernas y redes de conexión similares al tráfico ilícito de drogas (United Nations Environment Programme [UNEP], 2012), por ejemplo, el empleo de pistas de aterrizaje clandestinas que les permiten operar y conectarse entre países y continentes.

Como consecuencia, el alto nivel de organización de estos grupos criminales afecta sustancialmente las políticas públicas destinadas a resguardar

los activos ambientales, exigiendo el empleo y dedicación de muchos recursos humanos, financieros y económicos (European Union Agency for Law Enforcement Cooperation [EUROPOL], 2011). En este sentido, a medida que estos crímenes persisten, crecen y destruyen importantes ecosistemas terrestres, como la Amazonia y regiones fronterizas, los impactos ambientales son directos y visibles principalmente en áreas legalmente protegidas, convirtiéndolas en corredores aéreos, fluviales y terrestres para nuevas rutas y actividades vinculadas con el crimen organizado transnacional.

Por ejemplo, la región amazónica que se extiende a lo largo de varios países (Bolivia, Brasil, Colombia, Ecuador, Guayana, Guayana Francesa, Perú, Surinam y Venezuela) tiene un área aproximada de 8.470.209 km² equivalente al 40% del territorio de América del Sur (Red Amazónica de Información Socioambiental Georreferenciada [RAISG], 2020a), genera entre el 16% y 20% del agua dulce y el 10% de la biodiversidad mundial, produce más del 10% de oxígeno del planeta y es considerada la mayor selva tropical del mundo (Banco de Desarrollo de América Latina [CAF], 2019). Sin embargo, a pesar de ser una importante región estratégica debido a la abundancia de los recursos hídricos, energéticos y minerales que posee, es uno de los escenarios donde se verifican complejos CATs (Zabyelina et al., 2020; Lycia et al., 2019).

En los últimos 10 años la minería ilegal, el tráfico ilícito de drogas, la producción de cultivos ilícitos de coca (United Nations Office on Drugs and Crime [UNODC], 2012) y la construcción de pistas de aterrizaje clandestinas (Insight Crime, 2020; Ejército Brasileiro, 2018) se han intensificado (Lycia et al., 2019). Los cultivos ilícitos de coca, por ejemplo, a menudo están conectados y se superponen con áreas productoras de oro,

convirtiéndose en una estrategia por parte del crimen organizado, para diversificar sus fuentes de ingreso y al mismo tiempo permanecer activos en el tráfico ilícito de drogas (Zabyelina et al., 2020). Colombia, Perú y Bolivia han sido declarados por la UNODC, como las principales fuentes de producción de cultivos de coca y abastecimiento mundial de cocaína (UNODC, 2021a). Venezuela ha sido denunciada por tener participación y complicidad con el tráfico ilícito de drogas, minería ilegal y crímenes contra la humanidad (United Nations Human Rights Council [UNHRC], 2020; Insight Crime, 2020). Brasil registró un aumento de la deforestación del 34% en el año 2020 (Instituto Nacional de Pesquisas Espaciais [INPE], 2020) y varias investigaciones han identificado y confirmado la existencia de más de 2300 puntos y 245 áreas relacionadas con la minería ilegal de oro, diamantes y coltán, distribuidas por toda la región amazónica (RAISG, 2018a).

Por consiguiente, la presencia de grupos del crimen organizado en la selva amazónica ha convertido a la región suramericana en el epicentro mundial del comercio transnacional ilegal de recursos naturales y del tráfico ilícito de drogas (Global Initiative Against Transnational Organized Crime, 2016; UNODC, 2019c), generando como consecuencia una degradación ambiental acelerada que ha permitido abrir espacio a diferentes formas de violencia.

Organizaciones como la UNODC, la INTERPOL, la RAISG y el Centro de Gestión y Operación del Sistema de Protección de la Amazonia (CENSIPAM), han implementado el uso de técnicas de teledetección y extensas evaluaciones de campo, entre otras tecnologías geoespaciales, para el monitoreo y detección de delitos ambientales en la selva amazónica y regiones fronterizas; además de importantes contribuciones para la implementación de políticas públicas contra estos devastadores crímenes. Sin embargo, las metodologías aplicadas para la identificación y detección de las áreas afectadas se basan en expertos humanos, haciendo que las tareas de análisis e interpretación de las imágenes de satélite sean lentas y costosas.

Debido a su extensión, actuar en escenarios como la amazonia demanda métodos que permitan automatizar la detección de delitos ambientales a gran escala, para facilitar la planificación de operaciones de interdicción terrestre destinadas

a resguardar los activos ambientales de la región, además de proteger los recursos naturales a través de la formulación de normas legales nacionales e internacionales. En este contexto, la Inteligencia Geoespacial (*Geospatial Intelligence - GEOINT*) y la Inteligencia Artificial Geoespacial (*Geospatial Artificial Intelligence - GeoAI*), se convierten en importantes herramientas que permiten el análisis de datos geoespaciales para describir, evaluar y representar visualmente características y actividades georreferenciadas en la Tierra.

La GEOINT, se caracteriza por su capacidad de identificar, recopilar, almacenar y manipular datos para crear conocimiento geoespacial a través del pensamiento crítico, el razonamiento geoespacial y las técnicas analíticas; los datos pueden ser estructurados (información sobre ubicaciones y formas de características geográficas, como coordenadas geográficas, imágenes de satélite) y datos no estructurados (información geográfica que no está organizada de una manera predefinida, como texto con información geográfica de alguna actividad). Describe e interpreta el impacto humano de un determinado evento, permitiendo presentar éticamente el conocimiento para la toma de decisiones (Pennsylvania State University, 2020). Se diferencia de otras fuentes de recopilación de inteligencia, porque inherentemente integra y se enriquece de múltiples fuentes de inteligencia (*multiINT*) en un contexto espacio temporal (Clark, 2020).

En la actualidad la GEOINT, combina diferentes herramientas y tecnologías como la inteligencia artificial (*Artificial Intelligence - AI*), el aprendizaje automático (*Machine Learning - ML*), el *Big Data*, la experiencia humana para la recopilación de información y el procesamiento de datos procedentes de múltiples fuentes, destacándose el uso de fuentes abiertas (United States Geospatial Intelligence Foundation [USGIF], 2020; Clark, 2020; Coorey, 2018), las cuales no exigen ninguna restricción a su acceso y pueden ser obtenidas principalmente a través de internet (medios de comunicación digital, libros, artículos científicos, información geoespacial de libre acceso, entre otros) (Lowenthal, 2020), permitiendo adoptar nuevos enfoques para desarrollar una mayor conciencia situacional en la toma de decisiones.

En este sentido, su uso permite descubrir vínculos y patrones relevantes en los datos, para

proporcionar información y obtener ventajas en la toma de decisiones complejas (Letitia, 2013), por ejemplo, las destinadas a combatir el crimen organizado transnacional y los delitos ambientales.

La GeoAI es una subdisciplina de la inteligencia artificial que utiliza el aprendizaje automático, la ciencia de la información geográfica, la minería de datos y computación de alto rendimiento, para extraer y analizar información de grandes conjuntos de datos geoespaciales (United Nations Committee of Experts on Global Geospatial Information Management [UN-GGIM], 2020; VoPham et al., 2018). En este particular, el aprendizaje automático y en especial las redes neuronales profundas (*Deep Neural Networks – DNN*) han contribuido para resolver una gran variedad de problemas debido a su capacidad de aprender características representativas y discriminatorias a partir de los datos (Lecun et al., 2015).

Recientemente, en el marco de los programas de observación de la Tierra Copernicus de la Unión Europea y el Proyecto Landsat del Servicio Geológico de los Estados Unidos (USGS) y la Administración Nacional de Aeronáutica y del Espacio (NASA), varios satélites ópticos y de radar de apertura sintética se lanzaron con la finalidad de ofrecer a la comunidad científica grandes cantidades de datos de teledetección de libre acceso (Kussul et al., 2017). Esto ha promovido enfoques interdisciplinarios para estudiar la Tierra que han abierto oportunidades sin precedentes para una amplia gama de aplicaciones y crearon la base para el uso de métodos de ML (Vargas, 2020; Zhang et al., 2016) en muchas de las tareas de la comunidad de teledetección. Algunos enfoques se han dirigido a la clasificación de escenas, la detección y segmentación de objetos en imágenes de satélite (Janowicz et al., 2020), la fusión y el registro de imágenes (Ma et al., 2019), entre otros, con resultados favorables que han superado en algunos aspectos las capacidades de un operador humano.

Con el fin de fomentar el desarrollo de aplicaciones GeoAI para la clasificación de imágenes, destinadas a prevenir el impacto ambiental que producen las actividades del crimen organizado en la selva amazónica, es esencial disponer de grandes cantidades de datos correctamente etiquetados. Sin embargo, reunir esta gran cantidad de imágenes de satélite y recopilar información

asociada con las características reales del terreno (verdad de campo) para relacionar los datos de la imagen y entrenar modelos de DNN, es una tarea no trivial que requiere considerables esfuerzos de ingeniería y tiempo.

Actualmente, la Agencia Espacial Europea (ESA) a través del programa Copernicus, proporciona datos de observación de la Tierra de alta calidad, que se encuentran disponibles libremente con la intención de impulsar aplicaciones destinadas a mejorar la gestión ambiental, comprender y mitigar los efectos del cambio climático y garantizar la seguridad ciudadana (European Commission, 2015).

Considerando esta disponibilidad de grandes cantidades de datos de teledetección fiables y continuos, en este artículo se presenta una opción viable para la detección de CATs usando inteligencia artificial y métodos de inteligencia geoespacial para analizar imágenes de satélite y un banco de datos georreferenciados. El artículo ofrece tres principales contribuciones: 1) describe la metodología para la creación del conjunto de datos *AmazonCRIME*, compuesto por 30.000 imágenes multiespectrales de 13 bandas georreferenciadas, provenientes del satélite Sentinel-2, etiquetadas con 6 clases diferentes, destinadas a la identificación y clasificación de áreas vinculadas con CATs en la selva amazónica; 2) coloca a disposición de la comunidad científica el conjunto de datos propuesto; 3) demuestra a través de aplicaciones reales, el uso efectivo de *AmazonCRIME*, mediante el entrenamiento de modelos GeoAI cuyas predicciones identifican y clasifican CATs, en áreas naturales protegidas localizadas en la región amazónica.

2. Trabajos relacionados

El primer componente para un proyecto GeoAI está en la calidad, cantidad y obtención del conjunto de datos que se utilizará para el entrenamiento de los modelos DNN. Sin embargo, debido a las características propias de los datos de teledetección, existen desafíos en torno a la disponibilidad y la proporción de conjuntos de datos de entrenamiento etiquetados para extraer conocimientos de las imágenes de satélite (UN-GGIM, 2020), ya que generar estos conjuntos de datos es una tarea difícil y compleja (Hoeser et al., 2020). Para

superar estos obstáculos diversos investigadores han realizado estudios relacionados con la generación de conjuntos de datos de imágenes satelitales y fotografías aéreas, con la finalidad de contribuir con el desarrollo de aplicaciones específicas de GeoAI (Schmitt et al., 2021; Camps-Valls et al., 2021).

En este sentido en la Tabla 1, es presentada una breve descripción de trabajos relacionados con este dominio, donde es posible observar que,

para la detección de objetos y aplicaciones de segmentación semántica, los conjuntos de datos se componen principalmente de fotografías aéreas en color o imágenes de satélite de muy alta resolución espacial (alrededor de 50 cm). Al contrario, cuando se dispone de imágenes de menor resolución espacial, los conjuntos de datos se utilizan para tareas de clasificación, ya que la delimitación de objetos pequeños se hace difícil.

Tabla 1. Conjunto de datos para aplicaciones de GeoAI con imágenes de teledetección.

Conjunto de datos	Fuente	Tamaño de la imagen	Nº de imágenes	Tarea / Visión Computacional	Visión General / Aplicaciones	Referencia
DOTA	Fotografía aérea en color	800×800 a 4000×4000	2.806	Detección de objetos	15 clases 188.000 instancias	(Xia, et al., 2018)
Agriculture-Vision	Fotografía aérea en color	512×512	21.061	Segmentación de instancias	6 tipos de anotaciones	(Chiu, et al., 2020)
LandCover.ai	Fotografía aérea en color	9000×9000	41 ortofotos	Segmentación semántica	3 clases de cobertura del suelo	(Boguszews, et al., 2020)
UC Merced Land Use Dataset	Fotografía aérea en color	256×256	2.100	Clasificación	21 clases de cobertura del suelo	(Yang y Newsam, 2010)
SAT-4	Fotografía aérea en color	28 ×28	500.000	Clasificación	4 clases de cobertura del suelo	(Basu, et al., 2015)
SAT-6	Fotografía aérea en color	28×28	405.000	Clasificación	6 clases de cobertura del suelo	(Basu, et al., 2015)
Cactus Aerial Photos	Fotografía aérea en color	32×32	17.000	Clasificación	2 clases	(López, et al., 2019)
RESISC45	Fotografía aérea en color	256×256	31.500	Clasificación	45 clases	(Cheng y Lu, 2017)
DeepGlobe – Building Detection	WorldView-3 (0,3 m)	650×650	24.586	Segmentación semántica	1 clases	(Demir, et al., 2018)
Brazilian Cerrado-Savanna	RapidEye (0,3 m)	64×64	1.311	Clasificación	4 clases	(Nogueira, et al., 2016)
Brazilian Coffee Scenes Dataset	SPOT	64×64	51.004	Clasificación	2 clases	(Penatti y Dos Santos, 2015)
Planet	Planet’s Flock 2 (0,9 m)	256×256	150.000	Clasificación	13 categorías de cobertura del suelo / selva amazónica	(Planet Labs, 2017)
So2Sat LCZ42	Sentinel-1 Sentinel -2 (10 m)	32×32	400.000	Clasificación	17 clases	(Zhu, et al., 2020)
EuroSAT	Sentinel-2 (10 m)	64×64	27.000	Clasificación	10 clases de cobertura del suelo	(Helber y Borth, 2019)
BigEarthNet	Sentinel-2 (10 m)	120×120 60×60 20×20	590.326	Clasificación	43 clases de cobertura del suelo	(Sumbul y Markl, 2019)
OpenSARUrban	Sentinel -1 (10 m)	100×100	33.358	Clasificación	10 clases de interpretación urbana	(Zhao, et al., 2020)
SEN12MS	Sentinel-1, Sentinel-2 (10 m), MODIS Land Cover (250 m)	256×256	541.986	Clasificación	Cobertura del suelo MODIS pueden utilizarse como etiquetas o datos auxiliares	(Schmitt, et al., 2019)

Se observa que hay una gran variedad de conjuntos de datos especiales disponibles públicamente para fomentar la investigación relacionada con la GeoAI, se centran en aplicaciones agrícolas, el uso y la cobertura del suelo y la detección de características naturales y culturales comunes. Los datos de Sentinel se utilizan con frecuencia para componer este tipo de conjunto de datos, debido a que son gratuitos (de fácil acceso) y ofrecen una resolución espacial relativamente buena. Otro factor por considerar es la extensión del área de estudio. En el caso de grandes extensiones, el uso de imágenes de satélite ofrece ventajas con respecto a las fotografías aéreas.

3. Teledetección y Datos de Inteligencia Geoespacial

Se utilizaron datos disponibles libremente provenientes de fuentes *multi-INT* como OSINT (*Open-Source Intelligence*) y GEOINT (*Geospatial Intelligence*), para generar el conjunto de datos y obtener la verdad de campo. Inicialmente se hizo uso de la plataforma de *Google Earth Engine* (GEE) para descargar y procesar las imágenes multiespectrales Sentinel-2. Posteriormente para obtener la verdad de campo, se emplearon datos derivados de la RAISG, informes de monitoreo de territorios afectados por cultivos ilícitos de la UNODC, Agencia Nacional de Aviación Civil (ANAC), *openAIP*, MapBiomás Brasil y la Base de Datos Mundial sobre Áreas Protegidas (WDPA).

3.1. Sentinel-2

La misión Copernicus Sentinel-2 (European Space Agency [ESA], 2015) está conformada por una constelación de dos satélites en órbita polar heliosincrónica en fase de 180° entre sí. Su objetivo es monitorear la variabilidad de la superficie de la Tierra, captura imágenes en 13 bandas multiespectrales con resoluciones espaciales de 10 m, 20 m y 60 m. Posee un ancho de franja de hasta 290 km y una resolución temporal de 10 días en el Ecuador con un satélite y 5 días con dos satélites respectivamente; características que, en condiciones libre de nubes, permiten apoyar el monitoreo de la cobertura terrestre en tiempos oportunos. Los productos Sentinel-2 disponibles para los usuarios poseen niveles de procesamiento 1C y 2A. Para

el conjunto de datos *AmazonCRIME*, cada imagen está conformada por 13 bandas multiespectrales georreferenciadas, nivel 1C, con una cobertura de nubes máxima de 5%. Se generaron compuestos de imágenes con combinaciones de banda RGB (4,3,2) y color infrarrojo CIR (8, 4, 3). No se efectuó ningún procesamiento adicional.

3.2. Red Amazónica de Información Socioambiental Georreferenciada

La RAISG (RAISG, 2012) es un consorcio de organizaciones de la sociedad civil, conformado por seis países amazónicos (Bolivia, Brasil, Colombia, Ecuador, Perú y Venezuela), que, con apoyo de la cooperación internacional, tiene como principal objetivo generar y difundir mediante protocolos comunes, informes, datos estadísticos, sociales, ambientales y geoespaciales de la selva amazónica. Para obtener información de las áreas afectadas por la minería ilegal y deforestación, se descargaron datos del mapa de minería ilegal en Territorios Indígenas y Áreas Naturales Protegidas en la Amazonia (RAISG, 2018b) y de la sección de descargas de la página oficial (RAISG, 2020b).

3.3. Informes de Monitoreo de Territorios Afectados por Cultivos Ilícitos

El Programa de Monitoreo de Cultivos Ilícitos (ICMP) de la UNODC, actualmente promueve el desarrollo de un proyecto global que, a través de elementos técnicos y estratégicos centrados en evidencias técnicas y objetivas, apoya la realización de estudios de monitoreo sobre la coca en Bolivia, Colombia y Perú; sobre la amapola en Afganistán, México, Myanmar y un estudio sobre el cannabis en Nigeria (UNODC, 2021b; UNODC, 2021c). Proporciona apoyo técnico y supervisión directa en los diferentes sistemas nacionales de monitoreo con la finalidad de garantizar los estándares internacionales metodológicos de los datos que son producidos. Anualmente generan una serie de informes que, mediante el uso de técnicas de teledetección, sistemas de información geográfica (SIG) y extensas evaluaciones de campo, contemplan los resultados relacionados con la ubicación, extensión y evolución de las áreas que son afectadas por cultivos ilícitos. La verdad de campo de las áreas potenciales de cultivos ilícitos de coca para

el conjunto de datos *AmazonCRIME*, fue obtenida de los informes de monitoreo de territorio afectados por cultivos ilícitos de Bolivia (UNODC, 2017a; UNODC, 2019a), Colombia (UNODC, 2017b; UNODC, 2019b) y Perú (UNODC, 2017c; UNODC, 2018).

3.4. Agencia Nacional de Aviación Civil y OpenAIP

La ANAC (Republica Federativa do Brasil, 2005), es una institución federal encargada de regular y fiscalizar todas las actividades de aviación civil e infraestructura aeronáutica y aeroportuaria de Brasil. Entre sus variadas competencias, se destaca el registro y fiscalización de los aeródromos en el territorio nacional, así como también hacer disponibles datos que permitan promover investigaciones científicas y el desarrollo tecnológico. *OpenAIP (Open Aeronautical Information Publication)* (openAIP, 2021), por su parte, es una plataforma web, que ofrece información y datos aeronáuticos gratuitos y actualizados, entre los cuales se pueden obtener las coordenadas geográficas de los aeródromos que existen mundialmente. Los datos provenientes de estas fuentes de información permitieron obtener la ubicación de los aeródromos localizados en la selva amazónica.

3.5. MapBiomias Brasil

El proyecto MapBiomias (Souza et al., 2020), es una red de especialistas multidisciplinaria que, mediante el empleo de técnicas de teledetección, SIG, ciencias de la computación y procesamiento en la nube, genera series históricas anuales de mapas de uso y cobertura de la tierra de los diferentes biomas de Brasil, entre ellos el Amazonas, el cual es considerado uno de los biomas que tiene mayor presión sobre la cobertura terrestre original. Los datos geoespaciales generados por el proyecto MapBiomias, se utilizaron para identificar las áreas deforestadas, cuerpos de agua y bosques.

3.6. Base de Datos Mundial sobre Áreas Protegidas

La WDPA (Bingham et al., 2019), es un proyecto conjunto entre el Programa de las Naciones Unidas para el Medio Ambiente (PNUMA) y la Unión Internacional para la Conservación de la

Naturaleza (UICN), administrado por el Centro de Monitoreo de la Conservación Mundial del PNUMA (UNEP-WCMC), que, a través de su plataforma en línea, permite acceder y descargar libremente una gran variedad de información actualizada, relacionada con los datos mundiales sobre áreas protegidas terrestres y marinas. La selección de las áreas de bosque y cuerpos de agua del conjunto de datos *AmazonCRIME*, así como también la identificación de áreas protegidas para ejemplificar aplicaciones de uso real, se realizó con datos geoespaciales derivados de la WDPA.

4. Creación del conjunto de datos AmazonCRIME

Generar y etiquetar un conjunto de datos para clasificación de imágenes en combinación con métodos de GeoAI, requiere adquirir una gran cantidad de imágenes satelitales. Con el fin de garantizar la correcta identificación de la verdad de campo y automatizar el procesamiento, selección y descarga de las imágenes, se utilizó el software QGIS (QGIS Development Team, 2021), la plataforma de GEE (Gorelick et al., 2017), y la librería de código abierto para el procesamiento de datos geoespaciales GDAL (Warmerdam, 2008). A continuación, se describen cada una de las etapas.

4.1. Verdad de Campo

El conjunto de datos *AmazonCRIME*, está conformado por seis clases: Pistas de Aterrizaje (*Airstrips*), Deforestación (*Deforestation*), Bosque (*Forest*), Minería Illegal (*Illegal Mining*), Cultivos Ilícitos - Área Potencial de Cultivos de Coca (*Illicit Crops -Potential Coca Cultivation Area-PCCA*) y Agua (*Water*). Para seleccionar las muestras, se generaron datos geoespaciales (vectores) derivados de las diferentes fuentes de información y se visualizaron en conjunto con diferentes mosaicos de imágenes Sentinel-2 de la selva amazónica, construidos a través del complemento de GEE en QGIS. Los detalles individuales se describen en la Tabla 2.

4.2. Preparación de los datos en GEE

En esta etapa se realizó una adaptación de la metodología aplicada por (Schmitt et al., 2018), para seleccionar y procesar cada una de las

Tabla 2. Descripción de las clases que componen la verdad de campo de *AmazonCRIME*, fuente de datos y método de adquisición. El archivo vectorial de la Selva Amazónica de la RAISG se utilizó para delimitar el área de estudio.

Clase	Fuente de datos	Método de adquisición
Pistas de Aterrizaje (<i>Airstrips</i>)	-ANAC	-Identificación de 5.812 aeródromos de todo tipo (asfaltado, no asfaltado, grama, tierra, aeropuertos, entre otros), en los 9 países amazónicos.
	-openAIP	-Inspección visual para confirmar la ubicación y selección de los aeródromos ubicados en el área de estudio con características rurales. -Selección final de 2.481 aeródromos rurales. -Detección y selección de 187 nuevos aeródromos con características rudimentarias y aislados de centros poblados. -Selección final de 2.668 aeródromos.
Deforestación (<i>Deforestation</i>)	-RAISG	-Identificación y selección de 5.500 puntos representativos de esta clase.
	-MapBiomias	
Bosque (<i>Forest</i>)	-MapBiomias	-Identificación y selección de 5.500 puntos representativos de esta clase.
	-WDPA	
Minería Ilegal (<i>Illegal Mining</i>)	-RAISG	-Identificación y selección de 7.304 puntos representativos de esta clase.
Cultivos Ilícitos (<i>Illicit Crops_PCCA</i>)	-UNODC Bolivia	-Georreferenciación de los mapas de densidad de cultivos de coca generados en los informes de monitoreo de territorios afectados por cultivos ilícitos de:
	-UNODC Colombia	-Bolivia censos del año 2016 y 2018. Identificación y selección de 712 puntos.
	-UNODC Perú	-Colombia censos del año 2016 y 2018. Identificación y selección de 3.291 puntos. -Perú censos del año 2016 y 2017. Identificación y selección de 1.709 puntos. -Selección final de 5.712 puntos representativos de esta clase. -Los criterios considerados para la selección de las muestras se explican con mayor detalle en la preparación de los datos con GEE.
Agua (<i>Water</i>)	-MapBiomias -WDPA	-Identificación y selección de 5.500 puntos representativos de esta clase.

imágenes que estructuran el conjunto de datos. GEE almacena una gran cantidad de imágenes satelitales a escala global, que incluyen los datos generados por la misión Copernicus Sentinel-2. Proporciona herramientas y aplicaciones de programación que permiten procesar y analizar grandes conjuntos de datos geoespaciales (Gorelick et al., 2017). Actualmente ofrece productos de Sentinel-2 con niveles de procesamiento 1C y 2A (GEE, 2021). Los productos de nivel 1C, se encuentran ortorectificados en el sistema de referencia de coordenadas de datum WGS84, proyección UTM, con niveles digitales correspondientes a valores de reflectancia aparente por encima de la atmósfera (*Top-Of-Atmosphere reflectances – TOA*). Los productos de nivel 2A, son una evolución del nivel 1C en el que se aplican correcciones atmosféricas para proporcionar ortoimágenes con niveles digitales expresados en valores de reflectancia reales a nivel de superficie (*Bottom-Of-Atmosphere reflectances – BOA*). (ESA, 2015). En este sentido, esta plataforma se usó para obtener y exportar, recortes de imágenes de 256×256 píxeles, nivel

1C, en formato GeoTIFF, para cada clase. Se consideró la cobertura de nubes y la temporalidad de adquisición, como criterio para generar mosaicos de imágenes diferentes del área de estudio (región amazónica).

En este particular, es importante mencionar que GEE considera el criterio de cobertura de nubes en base a toda el área de interés (mosaico) y no por imagen individual, motivo por el cual no se utilizó un criterio de nubes generalizado para todos los rangos temporales. Este fue uno de los desafíos encontrados para generar mosaicos de imágenes con la menor cobertura de nubes posible, ya que, por las características climáticas propias de la región amazónica (altas temperaturas y frecuentes lluvias), resulta muy difícil generar mosaicos totalmente libres de nubes porque la región se mantiene nublada prácticamente durante todo el año.

En la Figura 1 se especifica el flujo de trabajo. A continuación, se describe en detalle cada una de las etapas:



Figura 1. Flujo de trabajo del método semiautomático para seleccionar y exportar los recortes de imágenes mediante GEE.

4.2.1. Pista de Aterrizaje

Fueron identificados y seleccionados 2.668 aeródromos, se generaron cuatro colecciones de imágenes con los siguientes criterios: para los años 2017 y 2018, con una cobertura de nubes del 1% y para los años 2016 y 2019, con una cobertura de nubes del 5%. Fueron exportados 2.668 recortes de imágenes por cada año, para un total de 10.672.

4.2.2. Minería Ilegal

Se generaron dos colecciones de imágenes, para el año 2017 con una cobertura de nubes del 5% y para el año 2019 con una cobertura de nubes del 1%. A partir de los vectores de la verdad de campo, se exportaron 3.804 muestras para el año 2017 y 3.500 muestras para el año 2019, obteniendo un total de 7.304 recortes de imágenes para esta clase.

4.2.3. Cultivos Ilícitos

Las plantaciones de cultivos de coca se caracterizan porque en un año presentan diferentes fenologías y se pueden obtener entre 4 y 5 cosechas (UNODC, 2019b), esta particularidad permite observar distintos comportamientos espectrales durante este periodo (baja/alta densidad foliar y baja/alta reflectancia del suelo) además de diferentes patrones y formas de los lotes de terreno que son empleados para este tipo de cultivo. A partir de estas variables, las siguientes etapas se realizaron para la identificación y selección de las muestras:

- Georreferenciación de los mapas de densidad de cultivos de coca publicados en los informes de monitoreo de territorio afectados por cultivos ilícitos de la UNODC.

- Selección de las muestras únicamente en las áreas de mayor densidad y territorios afectados permanentemente durante los últimos 10 años.
- Interpretación del comportamiento espectral en combinaciones RGB y CIR, tono, forma, textura, patrón y entorno geográfico. Estos elementos permitieron resaltar los objetos de interés (lotes de coca), discriminándolos de otras coberturas.
- Experiencia de campo considerando los criterios de la Primera Ley de la Geografía¹. Las áreas de cultivos ilícitos de coca se caracterizan por la presencia de lotes de terrenos (unidades espaciales) dedicados a este tipo de cultivo. Estos lotes son continuos geográficamente, se encuentran cercanos y tienden a ser más parecidos que los distantes. Poseen linderos identificables en el terreno, se encuentran plantados con el mismo tipo de especie (coca) y generalmente agrupan cultivos de la misma edad.
- Para los años correspondientes a los informes de monitoreo de cultivos ilícitos de la UNODC que fueron seleccionados, se generó una colección de imágenes por cada estación (verano, invierno, otoño, primavera), con una cobertura de nubes del 5% respectivamente. Para Bolivia y Colombia años 2016 y 2018, para Perú años 2016 y 2017.
- Muestras seleccionadas. Bolivia: 712, Colombia: 3.291, Perú: 1.709. Total: 5.712 muestras para esta clase.

¹ “Todos los lugares están relacionados entre sí, pero los lugares más próximos en el espacio tienen una relación mayor que los distantes” (Tobler, 1970).

4.2.4. Deforestación, Bosque y Agua

Se generaron dos colecciones de imágenes, para el año 2016 con una cobertura de nubes del 5% y para el año 2018 con una cobertura de nubes del 1%. A partir de los vectores de la verdad de campo, se exportaron 2.250 muestras para cada clase por año, obteniendo un total de 5.500 recortes de imágenes correspondientes por clase.

4.3. Exportación de las imágenes

Las imágenes seleccionadas por cada clase se exportaron para almacenamiento local en formato GeoTIFF, WGS-84 lat/long, nivel 1C, con una resolución espacial de 10 m, tamaño de 256×256 píxeles y 13 bandas multiespectrales respectivamente.

4.4. Inspección Visual

Posteriormente se generaron composiciones de imágenes RGB en formato JPG, para cada conjunto de imágenes/clase. En esta etapa se realizó una inspección visual con la finalidad de realizar la selección final de las mejores 5.000 muestras representativas por cada clase.

4.5. Estructura Final

Finalmente, se realizó la selección final de las muestras, las cuales permitieron generar el

conjunto de datos con imágenes de satélite denominado *AmazonCRIME*, conformado por los siguientes elementos: *AmazonCRIME_MS*, compuesto por 5.000 recortes de imágenes por cada clase, para un total de 30.000 recortes de imágenes seleccionados de 13 bandas multiespectrales – georreferenciadas en formato GeoTIFF, nivel 1C; el archivo *AmazonCRIME.csv*, con los metadatos del conjunto de datos (enumeración y etiqueta) y una versión derivada denominada *AmazonCRIME_RGB*, compuesto por las bandas 4,3,2 en formato JPG. El conjunto de datos *AmazonCRIME* en sus diferentes versiones, está disponible para el público en el repositorio: <https://github.com/jp-geoAI/AmazonCRIME.git>. Este documento debe citarse cuando el conjunto de datos sea utilizado con fines de investigación. En la Figura 2 se ilustra una visión general y algunos ejemplos.

5. Aplicaciones del conjunto de datos AmazonCRIME

Las redes neuronales convolucionales (*Convolutional Neural Networks - CNNs*) han demostrado su capacidad para resolver problemas de clasificación de imágenes utilizando modelos jerárquicos, millones de parámetros y grandes conjuntos de datos. En este contexto, se van a llevar a cabo dos experimentos: 1) verificar que



Figura 2. Algunas muestras de recortes de imágenes RGB por cada clase y descripción general del conjunto de datos propuesto, *AmazonCRIME*: 1) clase Pista de Aterrizaje; 2) clase Deforestación; 3) clase Bosque; 4) clase Minería Ilegal; 5) clase Cultivos ilícitos; 6) clase Agua.

el conjunto de datos es útil para la clasificación con un modelo CNNs entrenado desde cero; 2) aplicar al caso real de identificación de CATs en áreas naturales protegidas localizadas en la región amazónica. La arquitectura *DenseNet-201* (Huang et al., 2017), se seleccionó para el entrenamiento de los modelos, debido a sus capacidades para mejorar la precisión en tareas de clasificación de imágenes. Varios trabajos (Abdani et al., 2019; Koh et al., 2021; Khan et al., 2021) han demostrado que las CNNs pueden ser esencialmente más profundas, más precisas y eficientes de entrenar si contienen conexiones más cortas entre las capas cercanas a la entrada y las cercanas a la salida, este es el principio de la arquitectura *DenseNet-201*, conectar cada capa con todas las demás en forma de retroalimentación. Entre sus principales ventajas están el hecho de que este tipo de red reduce el problema de la desaparición de los gradientes, fortalece la propagación de características, promueve la reutilización de funciones y reduce considerablemente el número de parámetros de entrenamiento (Huang et al., 2017).

Se entrenaron dos modelos con la arquitectura *DenseNet-201*, un primer modelo con las combinaciones de banda RGB y un segundo modelo con las combinaciones de banda CIR. El entrenamiento de los modelos se realizó utilizando una unidad gráfica de procesamiento (*Graphics Processing Unit – GPU*) NVIDIA Tesla T4, mediante el empleo del lenguaje de programación Python versión 3.7 y la biblioteca de código abierto TensorFlow 2.

Para evaluar el modelo, se calculó la matriz de confusión (*Confusion Matrix*). Este cálculo permite evaluar el rendimiento de un modelo

de clasificación al contabilizar y visualizar los valores de las predicciones en comparación con los valores observados (reales), mostrando cuando una clase es confundida con otra. A partir de esta matriz pueden ser calculadas métricas que describen la calidad del producto (clasificación) como: 1) Exactitud Global (*Overall Accuracy*); 2) Precisión (*Precision*) y 3) Exhaustividad (*Recall*). La exactitud global permite medir el porcentaje de casos en que el modelo ha acertado en la clasificación, en relación con el total de los datos. La precisión, proporciona un valor de calidad relativo al número total de predicciones realizadas y la exhaustividad, proporciona un valor de calidad relativo al número total de muestra positivas. Para el entrenamiento de los modelos el conjunto de datos *AmazonCRIME* se dividió en tres subconjuntos: 1) 80% para datos de entrenamiento; 2) 10% para datos de validación; 3) 10% para datos de prueba. Los datos de entrenamiento son usados para entrenar el modelo. Los datos de validación son usados para ofrecer una evaluación del modelo que permite ajustar los hiperparámetros durante el entrenamiento. Los datos de prueba son usados para obtener una evaluación imparcial del modelo al final del entrenamiento, deben mantenerse separados del proceso de entrenamiento y son empleados apenas una vez para evaluar el rendimiento del modelo.

En la Tabla 3, se especifican las configuraciones del entrenamiento y los resultados obtenidos de las métricas de evaluación empleadas. En la Figura 3, se muestran las matrices de confusión de cada modelo.

Comparando los valores de las métricas de calidad de la clasificación, se puede apreciar que el

Tabla 3. Configuraciones del entrenamiento y resultados de las métricas de evaluación de los modelos entrenados con la arquitectura *DenseNet-201*, utilizando el conjunto de datos propuesto *AmazonCRIME*, con las combinaciones de banda RGB y CIR.

	DenseNet-201	
	Primer Modelo RGB	Segundo Modelo CIR
Conjunto de Datos (<i>Dataset</i>)	AmazonCRIME	
División (<i>Splitting</i>)	entrenamiento: 80%	validación: 10% prueba: 10%
Forma de Entrada (<i>Input shape</i>)	256×256×3	
Optimizador (<i>Optimizer</i>)	Adam	
Perdida (<i>Loss</i>)	Categorical Cross-Entropy	
Tamaño del Lote (<i>Batch size</i>)	64	
Exactitud Global (<i>Overall Accuracy</i>)	95,76%	96,56%
Precisión (<i>Precision</i>)	95,89%	96,66%
Exhaustividad (<i>Recall</i>)	95,73%	96,56%

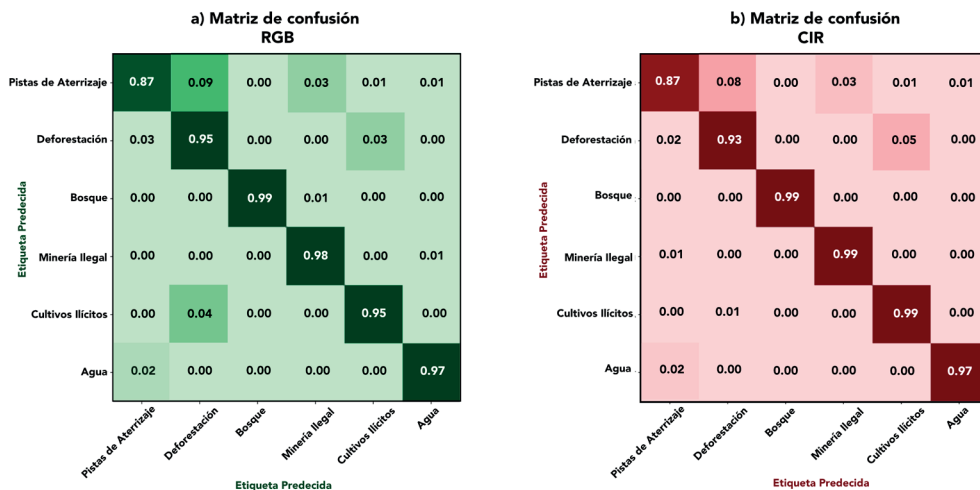


Figura 3. Matriz de confusión de los modelos entrenados con la arquitectura *DenseNet-201*, utilizando el conjunto de datos propuesto *AmazonCRIME*: a) matriz de confusión con los conjuntos de datos en combinaciones de banda RGB; b) matriz de confusión con los conjuntos de datos en combinaciones de banda CIR.

segundo modelo presentó un rendimiento superior en comparación con el primer modelo. Estos resultados comprueban la importancia de incluir información espectral en el infrarrojo próximo en los estudios de monitoreo ambiental y detección de alteraciones de la cobertura de la tierra en la región amazónica. La composición de bandas CIR muestra mayor sensibilidad a la discriminación de biomasa y detección del vigor de la vegetación en la región del espectro electromagnético del rango del infrarrojo, bien como la fuerte banda de absorción causada por la presencia de clorofila en la región visible, especialmente en la banda del rojo. Al respecto, la composición de bandas CIR, permitió una extracción mayor de características durante el entrenamiento, ya que los objetos de interés que representan cada clase del conjunto de datos se encuentran predominantemente influenciadas por la alta presencia de vegetación en el contexto espacial.

Se observó que ambos modelos tienden a confundir las clases “pistas de aterrizaje” y “cultivos ilícitos” con la clase “deforestación”, probablemente debido a que el elemento en común entre estas clases es la ausencia de cobertura forestal y la exposición de suelo desnudo en superficie, y en algunos de los recortes de imágenes del conjunto de datos, es prácticamente inevitable que se encuentren en una misma imagen la presencia de estas tres clases. En la Figura 4, se ejemplifican

algunas de las imágenes que el modelo confundió al clasificarlas.

Algo similar ocurre con la clase de “minería ilegal”, al confundirse con la clase “deforestación”, y con la clase “agua”, ya que las actividades de minería en la mayoría de las veces son ejercidas en las proximidades de las redes fluviales o cuerpos de agua, generan fuertes degradaciones de la cobertura vegetal y presentan elementos pictomorfológicos similares (tono, textura, patrón, entre otros). En la Figura 5 se presentan algunos ejemplos.

Posteriormente, para demostrar aplicaciones de uso real y probar la capacidad de generalización de los modelos entrenados, seleccionamos tres áreas naturales protegidas de la región amazónica, que presentan fuertes impactos ambientales por parte de grupos del crimen organizado; Parque Nacional Cerro Yapacana (Organización de los Estados Americanos [OEA], 2007), Parque Nacional Natural La Paya (MinAmbiente, 1978) y *Floresta Nacional do Amaná (Instituto Chico Mendes de Conservação da Biodiversidade [ICMBio], 2016)*.

Siguiendo la metodología aplicada en la sección 4.2, se generaron tres nuevas colecciones de imágenes Sentinel-2 (multiespectrales, georreferenciadas en formato GeoTIFF) con su respectiva enumeración por cada clase, desde el 01/01/2020 hasta el 31/07/2020, correspondientes a cada área

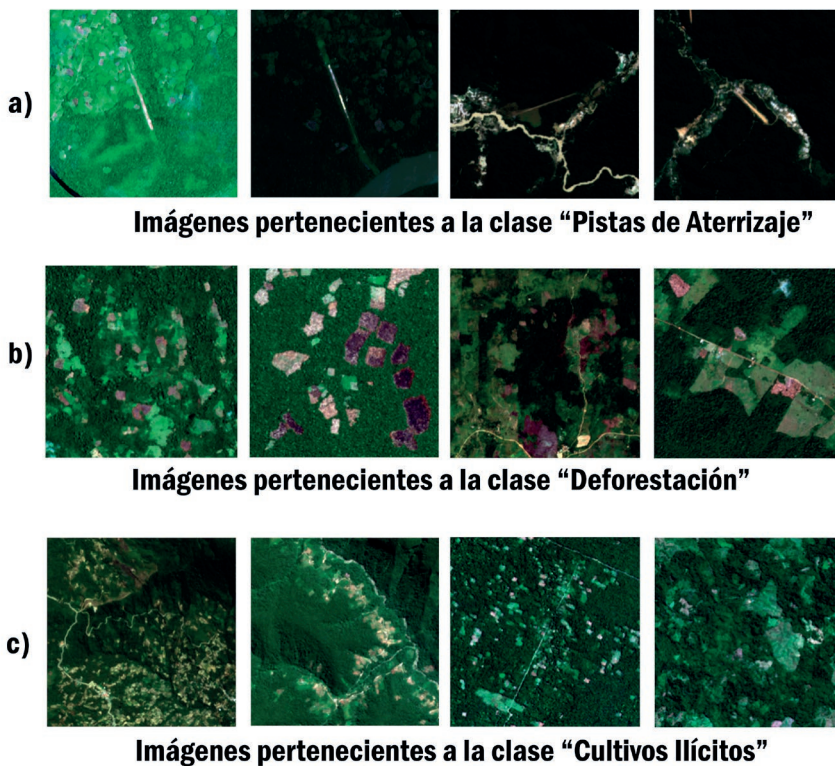


Figura 4. Ejemplo de imágenes clasificadas incorrectamente que tienen en común la clase deforestación. a) algunas de las imágenes pertenecientes a la clase pistas de aterrizaje, presentan en su contexto geográfico pérdida de la cobertura vegetal por causa de la deforestación, cultivos ilícitos y la minería ilegal; b) imágenes pertenecientes a la clase deforestación clasificadas incorrectamente como cultivos ilícitos; c) algunas de las imágenes pertenecientes a la clase cultivos ilícitos que fueron clasificadas incorrectamente ya que poseen también unidades y patrones espaciales correspondientes a la deforestación.



Figura 5. Ejemplo de imágenes clasificadas incorrectamente que tienen en común la clase agua. a) algunas de las imágenes pertenecientes a la clase minería ilegal que por las características propias de esta actividad que es ejercida cerca de las redes fluviales, se clasificaron incorrectamente como la clase agua; b) imágenes pertenecientes a la clase agua clasificadas correctamente.

natural protegida seleccionada, con una cobertura de nubes del 5%. El área del Parque Nacional Yapacana se dividió en 464 recortes de imágenes, El Parque Nacional Natural La Paya en 802 recortes de imágenes y la *Floresta Nacional do Amana* en 1.114 recortes de imágenes, todos los recortes con dimensiones de 256×256 píxeles. Posteriormente se exportaron para almacenamiento local y se derivaron compuestos enumerados de imágenes RGB y CIR en formato JPG.

Seguidamente, los modelos entrenados se aplicaron a los conjuntos de imágenes RGB y CIR (formato JPG) que fueron derivados, obteniendo predicciones que permitieron reconocer diversas áreas afectadas por los CATs, así como también identificar la imagen georreferenciada (formato GeoTIFF) equivalente con la enumeración y obtener las coordenadas geográficas, ya que en la realidad, este es el dato de mayor interés que se desea conocer para promover acciones destinadas a mitigar el impacto ambiental y planificar

Tabla 4. Resultados de las métricas de evaluación del proceso de inferencia, utilizando los conjuntos de imágenes RGB y CIR (formato JPG) que se derivaron.

	DenseNet-201 - Inferencia	
	Primer Modelo RGB	Segundo Modelo CIR
Exactitud Global (<i>Overall Accuracy</i>)	91,74%	93,68%
Precisión (<i>Precision</i>)	91,85%	93,77%
Exhaustividad (<i>Recall</i>)	91,69%	93,68%

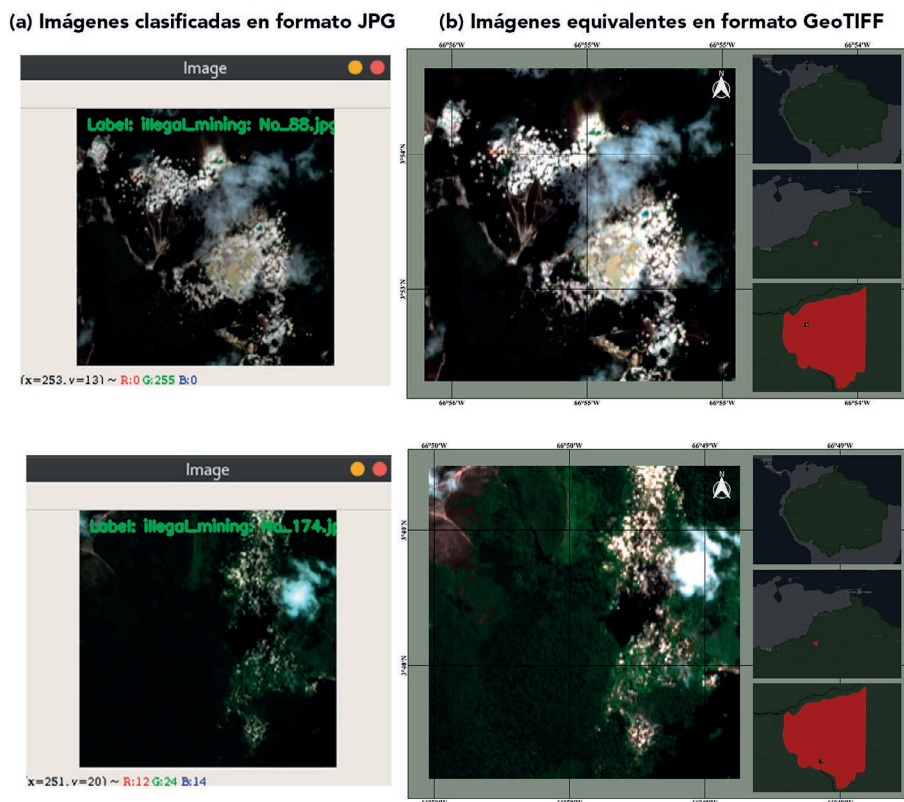


Figura 6. Clasificación y reconocimiento de áreas afectadas por la minería ilegal en el Parque Nacional Yapacana – Venezuela. a) ejemplos de imágenes en formato JPG con su respectiva etiqueta y enumeración clasificadas correctamente como “minería ilegal” durante el proceso de inferencia; b) representación de las imágenes georreferenciadas equivalentes de las imágenes JPG clasificadas. Los mapas de localización representan cartográficamente la región amazónica en América del Sur (polígono verde) y el Parque Nacional Yapacana (polígono rojo) en el estado Amazonas de Venezuela.

operaciones de interdicción terrestre, destinadas a combatir las actividades vinculadas con los CATs y el tráfico ilícito de drogas. Es importante mencionar que *AmazonCRIME* fue construido con imágenes de 2016 a 2019, por lo tanto, este conjunto de imágenes es totalmente desconocido por la red previamente entrenada. En la Tabla 4 se especifican los resultados obtenidos durante el proceso de inferencia.

En la Figura 6, se puede apreciar la clasificación y reconocimiento de áreas afectadas por la minería ilegal en el Parque Nacional Yapacana, este se encuentra localizado en el estado Amazonas de Venezuela, tiene un área aproximada de 320.000 ha, con un relieve en forma de meseta, característico de los tepuyes de la región amazónica venezolana (OEA, 2007). Actualmente se

encuentra fuertemente amenazado por la minería ilegal, que es ejercida por grupos armados que han encontrado en este ecosistema una fuente económica para financiar sus actividades vinculadas con el crimen organizado transnacional (RAISG, 2019).

Otro ejemplo se puede ilustrar en la Figura 7, al aplicar el modelo entrenado se logró obtener predicciones que permitieron reconocer varias áreas potenciales de cultivos ilícitos en el Parque Nacional Natural La Paya. Se encuentra localizado en el departamento del Putumayo al sur de Colombia, posee un área de 422.000 ha, se caracteriza por un sistema hídrico denso y una variada vegetación húmeda tropical. Entre las actividades antrópicas más frecuentes se destaca la extracción de madera, pesca, minería ilegal y cultivos ilícitos (MinAmbiente, 1978; Ministerio de Defensa,

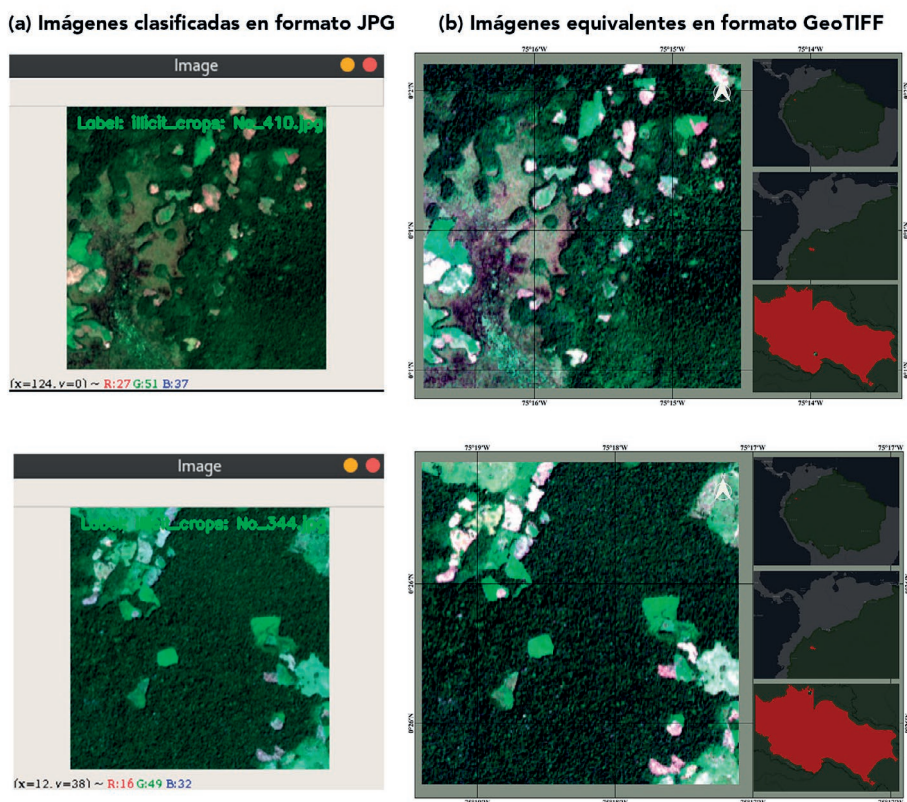
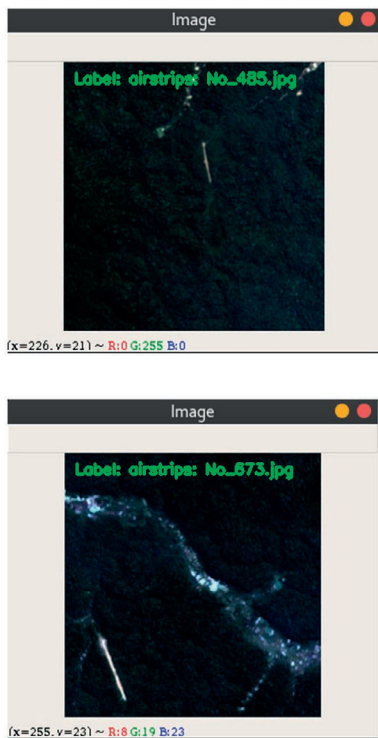


Figura 7. Clasificación y reconocimiento de áreas potenciales de cultivos ilícitos en el Parque Nacional Natural La Paya – Colombia. a) ejemplos de imágenes en formato JPG con su respectiva etiqueta y enumeración clasificadas correctamente como “cultivos ilícitos” durante el proceso de inferencia; b) representación de las imágenes georreferenciadas equivalentes de las imágenes JPG clasificadas. Los mapas de localización representan cartográficamente la región amazónica en América del Sur (polígono verde) y el Parque Nacional Natural La Paya (polígono rojo) en el departamento del Putumayo, Colombia.

(a) Imágenes clasificadas en formato JPG



(b) Imágenes equivalentes en formato GeoTIFF

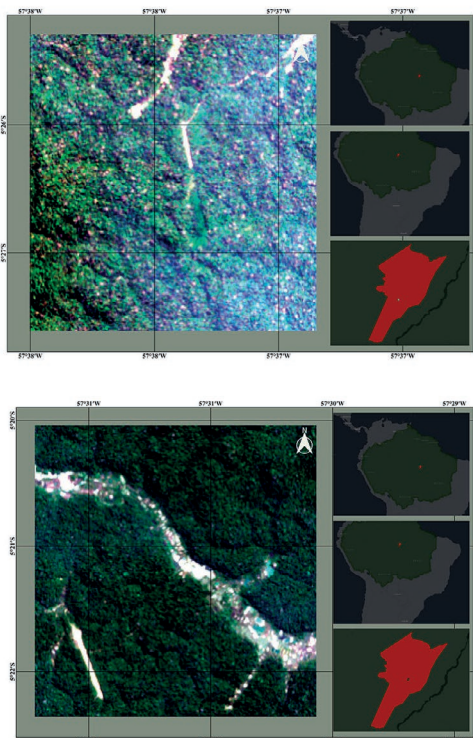


Figura 8. Clasificación y reconocimiento de pistas de aterrizaje en La Floresta Nacional do Amana – Brasil. a) ejemplos de imágenes en formato JPG con su respectiva etiqueta y enumeración clasificadas correctamente como “pistas de aterrizaje” durante el proceso de inferencia; b) representación de las imágenes georreferenciadas equivalentes de las imágenes JPG clasificadas. Los mapas de localización representan cartográficamente la región amazónica en América del Sur (polígono verde) y La Floresta Nacional do Amana (polígono rojo) en el estado de Para, Brasil.

2020), siendo esta área protegida una de las más afectadas por la deforestación ocasionada por grupos armados ilegales dedicados al tráfico ilícito de drogas.

La Floresta Nacional do Amana se encuentra en el estado de Para, Brasil, cuenta con un área aproximada de 540.417 ha, se caracteriza por tener una vegetación densa, extensa hidrografía y una variada diversidad de flora y fauna. Se creó con el objeto de ser un área protegida con uso sostenible de recursos naturales, sin embargo, la expansión de la minería ilegal y la deforestación, son las principales amenazas en esta unidad de conservación ambiental (ICMBio., 2016; Oliveira, 2018). En este sentido, al aplicar el modelo entrenado, las predicciones lograron identificar áreas afectadas por la minería de recursos naturales, así como también pistas de aterrizaje con características

rudimentarias. En este particular, al superponer los resultados con los datos geoespaciales de la ANAC, se logró identificar una pista de aterrizaje que no coincide con los registros oficiales. En la Figura 8, se muestran algunos ejemplos.

6. Conclusiones

En este documento abordamos el desafío de clasificar imágenes en la selva amazónica. Presentamos un nuevo conjunto de datos construido a partir de imágenes satelitales Sentinel-2, las cuales son de libre acceso y proporcionadas por el programa Copernicus de la Unión Europea. El conjunto de datos consta de 6 clases, 5.000 imágenes por clase, para un total de 30.000 imágenes multiespectrales de 13 bandas, nivel 1C, etiquetadas, enumeradas y georreferenciadas en formato

GeoTIFF. Adicionalmente ofrecemos una versión derivada, con combinaciones de banda RGB en formato JPG, empleada para entrenar la arquitectura *DenseNet-201*. Entre las principales contribuciones de este trabajo, presentamos el primer conjunto de datos destinado a generar aplicaciones de inteligencia artificial geoespacial, para la clasificación y detección de áreas vinculadas con crímenes ambientales transnacionales en la selva amazónica a gran escala, ya que de acuerdo a la revisión de los trabajos relacionadas, los conjuntos de datos disponibles están dirigidos principalmente para aplicaciones agrícolas, uso y cobertura del suelo y detección de elementos naturales y culturales comunes, siendo escasa la disponibilidad de conjuntos de datos de referencia para aplicaciones del dominio de interés de esta investigación. Destacamos la importancia de crear nuevos conjuntos de datos que prioricen la representación geográfica de la región amazónica, ya que la mayoría de los conjuntos de datos de imágenes de satélite disponibles públicamente, están sesgados hacia otras regiones del mundo, lo que limita la aplicación de la GeoAI en la Amazonia. Entrenamos y analizamos el rendimiento de las bandas espectrales RGB (4,3,2) y CIR (8,4,3), basado en recortes de imágenes y tareas de clasificación de escenas. Como resultado la combinación de bandas CIR obtuvo un rendimiento general superior en la calidad de los resultados, en comparación con la combinación de bandas RGB. Demostramos la capacidad de generalización de los modelos entrenados con el conjunto de datos *AmazonCRIME* y el potencial de uso para aplicaciones reales, al ser aplicados en nuevas imágenes de satélite, correspondientes a áreas de la región amazónica fuertemente amenazadas por el crimen organizado transnacional.

Esperamos que *AmazonCRIME* fomente el desarrollo de futuras investigaciones que permitan promover aplicaciones e iniciativas tanto de alcance local como regional, destinadas a resguardar los activos ambientales de tan importante región estratégica. La fuerte relación entre la minería ilegal, la deforestación con otras economías ilegales como el tráfico ilícito de drogas, se convierte en una amenaza que requiere de acciones urgentes, de enfoques integrales y de herramientas tecnológicas como la GeoAI, para restablecer el equilibrio ecológico en la selva tropical más grande del mundo.

7. Agradecimientos

Agradecemos al Programa de Posgraduación en Ciencias Geodésicas de la Universidad Federal de Paraná y el apoyo financiero al Consejo Nacional de Desarrollo Científico y Tecnológico de Brasil (CNPq) (190032/2017-0).

8. Referencias

- Abdani, S.R., & Zulkifley, M.A. 2019. Densenet with spatial pyramid pooling for industrial oil palm plantation detection. *In 2019 International Conference on Mechatronics, Robotics and Systems Engineering*. 134-138. <https://doi.org/10.1109/MoRSE48060.2019.8998735>
- Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., & Nemani, R. 2015. DeepSat - A Learning framework for Satellite Imagery. *Association for Computing Machinery*, 37, 1-10. <https://doi.org/10.1145/2820783.2820816>
- Bingham, H., Bignoli, D., Lewis, E., MacSharry, B., Burgess, N., Visconti, P., Kingston, N. 2019. Sixty years of tracking conservation progress using the World Database on Protected Areas. *Nature Ecology & Evolution*, 3, 737-743. <https://doi.org/10.1038/s41559-019-0869-3>
- Boguszewski, A., Batorski, D., Jankowska, N., Zambrzycka, A., & Dziedzic, T. 2020. LandCover.ai: Dataset for Automatic Mapping of Buildings, Woodlands and Water from Aerial Imagery. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1102-1110. <https://doi.org/10.1109/CVPRW53098.2021.00121>
- CAF (Banco de Desarrollo de América Latina). 2019. *La riqueza natural de la Amazonia como base del desarrollo sostenible regional*. Último acceso: 01 de Noviembre, 2021, de <https://www.caf.com/es/conocimiento/visiones/2019/09/la-riqueza-natural-de-la-amazonia-como-base-del-desarrollo-sostenible-regional/>.
- Camps-Valls, G., Tuia, D., Zhu, X.X., & Reichstein, M. (Eds.). 2021. *Deep learning for the Earth Sciences: A comprehensive approach to remote sensing, climate science and geosciences*. John Wiley & Sons. <https://doi.org/10.1002/9781119646181>
- Cheng, G., Han, J., & Lu, X. 2017. Remote Sensing Image Scene Classification: Benchmark and State of the Art. *Proceedings of the IEEE*, 105(10), 1865-1883. <https://doi.org/10.1109/JPROC.2017.2675998>

- Chiu, M.T., Xu, X., Wei, Y., Huang, Z., Schwing, A. G., Brunner, R.,... & Shi, H. 2020. Agriculture-vision: A large aerial image database for agricultural pattern analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2828-2838. <https://doi.org/10.1109/CVPR42600.2020.00290>
- Clark, R. 2020. *Geospatial Intelligence. Origins and Evolution*. Washington, DC: Geogertown University Press.
- Coorey, R. 2018. *The Evolution of Geospatial Intelligence. Australian Contributions to Strategic and Military Geography. Advances in Military Geosciences*. Springer, Cham. https://doi.org/10.1007/978-3-319-73408-8_10
- Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S.,... & Raskar, R. 2018. Deepglobe 2018: A challenge to parse the earth through satellite images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 172-181. <https://doi.org/10.1109/CVPRW.2018.00031>
- ESA (European Space Agency) 2015. *Sentinel -2 User Handbook*. ESA Standard Document.
- European Commission. 2015. *Copernicus. Europe's eyes on Earth*. Brussels: Publications Office of the European Union.
- EUROPOL (European Union Agency for Law Enforcement Cooperation). 2011. *EU Organised Crime Threat Assessment. OCTA 2011*. EUROPOL Public Information.
- Exército Brasileiro. 2018. *Operação Curare IX - Interdição de pista de pouso clandestina*. Último acceso: 16 de Febrero, 2021, de http://www.eb.mil.br/web/noticias/noticiario-do-exercito/-/asset_publisher/MjaG93KcunQI/content/no-contexto-da-operacao-curare-ix-1-brigada-infantaria-de-selva-realiza-interdicao-de-pista-de-pouso-clandestina-/8357041.
- GEE (Google Earth Engine). *Sentinel-2*. Último acceso 05/11/2021, de <https://developers.google.com/earth-engine/datasets/catalog/sentinel-2>.
- Global Initiative Against Transnational Organized Crime. 2016. *Organized Crime and Illegally Mined Gold in Latin America*. <https://globalinitiative.net/analysis/organized-crime-and-illegally-mined-gold-in-latin-america>
- Global Initiative Against Transnational Organized Crime. 2021. *Environmental crime: The not-so-hidden obstacle to combat climate change*. Último acceso: 01 de Noviembre, 2021, de <https://globalinitiative.net/analysis/environmental-crime-climate-change/>.
- Gore, M.L., Braszak, P., Brown, J., Cassey, P., Duffy, R., Fisher, J.,... & White, R. 2019. Transnational environmental crime threatens sustainable development. *Nature Sustainability*, 2(9), 784-786. <https://doi.org/10.1038/s41893-019-0363-6>
- 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>
- Helber, P., Bischke, B., Dengel, A., & Borth, D. 2019. Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(7), 2217-2226. <https://doi.org/10.1109/JSTARS.2019.2918242>
- Hoeser, T., & Kuenzer, C. 2020. Object detection and image segmentation with deep learning on earth observation data: A review-part i: Evolution and recent trends. *Remote Sensing*, 12(10), 1667. <https://doi.org/10.3390/rs12101667>
- Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. 2017. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 4700-4708. <https://doi.org/10.1109/CVPR.2017.243>
- ICMBio. 2016. Plano de Manejo Floresta Nacional do Amaná. Último acceso: 10 de Marzo, 2021, de <https://www.icmbio.gov.br/portal/unidadesdeconservacao/biomas-brasileiros/amazonia/unidades-de-conservacao-amazonia/1955-flona-do-amana>.
- INPE (Instituto Nacional de Pesquisas Espaciais). 2020. *TerraBrasilis*. Último acceso: 11 de Enero, 2021, de <http://terrabrasilis.dpi.inpe.br/en/home-page/>.
- Insight Crime. 2020. *Narcovuelos y pistas clandestinas en al menos seis estados de Venezuela*. Último acceso: 11 de Enero, 2021, de <https://es.insightcrime.org/noticias/noticias-del-dia/narcovuelos-seis-estados-venezuela/>.
- INTERPOL (International Criminal Police Organization). 2018. *World atlas of illicit flows*. Último acceso: 01 de Noviembre, 2021, de <https://globalinitiative.net/wp-content/uploads/2018/09/Atlas-Ilicit-Flows-FINAL-WEB-VERSION-copia-compressed.pdf>.
- Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. 2020. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 34(4), 625-636. <https://doi.org/10.1080/13658816.2019.1684500>

- Khan, M.A., Hussain, N., Majid, A., Alhaisoni, M., Syed Ahmad Chan, B., Kadry, S., Yu-Dong, Z. 2021. Classification of positive COVID-19 CT scans using deep learning. *Computers, Materials, & Continua*, 66(3), 2923-2938. <https://doi.org/10.32604/cmc.2021.013191>
- Koh, J. C., Spangenberg, G., & Kant, S. 2021. Automated Machine Learning for High-Throughput Image-Based Plant Phenotyping. *Remote Sensing*, 13(5), 858. <https://doi.org/10.3390/rs13050858>
- Kussul, N., Lavreniuk, M., Skakun, S., & Shelestov, A. 2017. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, 14(5), 778-782. <https://doi.org/10.1109/LGRS.2017.2681128>
- LeCun, Y., Bengio, Y., & Hinton, G. 2015. Deep learning. *nature*, 521(7553), 436-444. <https://doi.org/10.1038/nature14539>
- Letitia, A. 2013. Activity based intelligence: Understanding the unknown. *The intelligencer: Journal of US intelligence studies*, 20(2), 7-16.
- López-Jiménez, E., Vasquez-Gomez, J. I., Sanchez-Acevedo, M. A., Herrera-Lozada, J. C., & Uriarte-Arcia, A. V. 2019. Columnar cactus recognition in aerial images using a deep learning approach. *Ecological Informatics*, 52, 131-138. <https://doi.org/10.1016/j.ecoinf.2019.05.005>
- Lowenthal, M. 2020. *Intelligence. From secrets to policy*. CQ. Press.
- Lycia, B., Abdenur, A., Pellegrino, A., Porto, C., & Brasil, L. 2019. *Los delitos Ambientales en la Cuenca del Amazonas: el rol del crimen organizado en la minería*. El Pacto. Europa Latinoamerica. Programa de Asistencia contra el Crimen Transnacional Organizado. Último acceso: 11 de Mayo, 2020, de <https://www.elpaccto.eu/wp-content/uploads/2019/05/Los-Delitos-Ambientales-en-la-Cuenca-del-Amazonas-comprimido.pdf>.
- Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 166-177. <https://doi.org/10.1016/j.isprsjprs.2019.04.015>
- MinAmbiente. Parques Nacionales Naturales de Colombia. 1978. Plan de Manejo del Parque Nacional Natural La Paya. Último acceso: 10 de Marzo, 2021, de <https://www.parquesnacionales.gov.co/portal/wp-content/uploads/2020/10/plan-de-manejo-pnn-la-paya.pdf>.
- Ministerio de Defensa. 2020. *Medio Ambiente. Defensa*. Último acceso: 10 de Marzo, 2021, de https://www.mindefensa.gov.co/irj/go/km/docs/Mindefensa/Documentos/descargas/Documentos_Descargables/espanol/Medio%20Ambiente.pdf.
- Nogueira, K., Dos Santos, J.A., Fornazari, T., Silva, T.S.F., Morellato, L.P., & Torres, R.D.S. 2016. Towards vegetation species discrimination by using data-driven descriptors. In *2016 9th IAPR Workshop on Pattern Recognition in Remote Sensing (PRRS)* (pp. 1-6). IEEE. <https://doi.org/10.1109/PRRS.2016.7867024>
- OEA (Organización de los Estados Americanos). 2007. *Yapacana National Park*. BioParques: Parkswatch Program. Último acceso: 10 de Marzo, 2021, de [https://www.oas.org/dsd/AAPAD2/Docs/\(iii\)%20Yapacana%20NP%20Special%20Report%20\(Venezuela\).pdf](https://www.oas.org/dsd/AAPAD2/Docs/(iii)%20Yapacana%20NP%20Special%20Report%20(Venezuela).pdf).
- Oliveira, D. 2018. *Atividade garimpeira na região do Tapajós (PA): o caso na Flona do Amaná*. Monografía, Instituto CEUB de Pesquisa e Desenvolvimento, Centro Universitário de Brasília). <https://repositorio.uniceub.br/jspui/handle/235/11514>.
- openAIP. 2021. *Worldwide aviation database*. Último acceso: 5 de Enero, 2021, de <http://www.openaip.net/>.
- Penatti, O.A., Nogueira, K., & Dos Santos, J.A. 2015. Do deep features generalize from everyday objects to remote sensing and aerial scenes domains? In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 44-51. <https://doi.org/10.1109/CVPRW.2015.7301382>
- Pennsylvania State University. 2020. GEOINT MOOC. Último acceso 05/11/2021, de <https://www.e-education.psu.edu/geointmooc/node/1989>.
- Planet Labs. 2017. *Planet: Understanding the Amazon from Space*. Último acceso: 08 de Enero, 2021, de <https://www.kaggle.com/c/planet-understanding-the-amazon-from-space>.
- QGIS Development Team. *QGIS. A Free and Open Source Geographic Information System*. Último acceso: 08 de Enero, 2021, de <https://qgis.org/en/site/>.
- RAISG (Red Amazónica de Información Socioambiental Georreferenciada). 2012. *Amazonia Bajo Presión*. Último acceso: 14 de Enero, 2021, de www.raisg.socioambiental.org.
- RAISG (Red Amazónica de Información Socioambiental Georreferenciada). 2018a. *Amazonia saqueada: el primer mapa de minería ilegal en el pulmón del mundo*. 2018. Último acceso: 14 de Enero, 2021, de <https://www.amazoniasocioambiental.org/es/radar/mapa-inedito-indica-epidemia-de-garimpo-ilegal-na-panamazonia/>.

- RAISG (Red Amazónica de Información Socioambiental Georreferenciada). 2018b. *Presiones y Amenazas*. Último acceso: 14 de Enero, 2021, de <https://www.amazoniasocioambiental.org/es/mapas/#!/presiones>.
- RAISG (Red Amazónica de Información Socioambiental Georreferenciada). 2019. *¡Imperdonable! Parque nacional en Amazonas es devastado por la minería ilegal que dirige el ELN*. Último acceso: 14 de Enero, 2021, de <https://www.amazoniasocioambiental.org/es/radar/imperdonable-parque-nacional-en-amazonas-es-devastado-por-la-mineria-ilegal-que-dirige-el-eln/>.
- RAISG (Red Amazónica de Información Socioambiental Georreferenciada). 2020a. *Atlas Amazonia Bajo Presión 2020*. Último acceso: 01 de Noviembre, 2021, de <https://www.amazoniasocioambiental.org/es/publicacion/amazonia-bajo-presion-2020/>.
- RAISG (Red Amazónica de Información Socioambiental Georreferenciada). 2020b. *Cartographic Data*. Último acceso: 14 de Enero, 2021, de <https://www.amazoniasocioambiental.org/en/maps/#download>.
- Republica Federativa do Brasil. 2005. *LEI Nº 11.182, DE 27 DE SETEMBRO DE 2005*. (Cria a Agência Nacional de Aviação Civil – ANAC) Último acceso: 20 de Febrero, 2021, de http://www.planalto.gov.br/ccivil_03/_ato2004-2006/2005/Lei/L11182.htm.
- Schmitt, M., & Wu, Y. L. 2021. Remote Sensing Image Classification with the SEN12MS Dataset. *arXiv preprint arXiv:2104.00704*.
- Schmitt, M., Hughes, L., & Zhu, X. 2018. The SEN1-2 Dataset for Deep Learning in SAR-Optical Data Fusion. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-1*, 141-146. <https://doi.org/10.5194/isprs-annals-IV-1-141-2018>
- Schmitt, M., Hughes, L., Qiu, C., & Zhu, X. 2019. SEN12MS - A Curated Dataset of Georeferenced Multi-Spectral Sentinel-1/2 Imagery for Deep Learning and Data Fusion. *SPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W7*, 153-160. <https://doi.org/10.5194/isprs-annals-IV-2-W7-153-2019>
- Souza, C., Shimbo, J., Rosa, M., Parente, L., Alencar, A., Rudorff, B., Azevedo, T. 2020. Reconstructing Three Decades of Land Use and Land Cover Changes in Brazilian Biomes with Landsat Archive and Earth Engine. *Remote Sensing, 12*(17), 2735. <https://doi.org/10.3390/rs12172735>
- Sumbul, G., Charfuelan, M., Demir, B., & Markl, V. 2019. Bigearthnet: A Large-Scale Benchmark Archive for Remote Sensing Image Understanding. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 5901-5904. <https://doi.org/10.1109/IGARSS.2019.8900532>
- Tobler, W. R. 1970. A computer movie simulating urban growth in the Detroit region. *Economic geography, 46*(sup1), 234-240. <https://doi.org/10.2307/143141>
- UNEP (United Nations Environment Programme). 2012. *Transnational Environmental Crime - a common crime in need of better enforcement*. Último acceso: 22 de Febrero, 2021, de <https://wedocs.unep.org/rest/bitstreams/14319/retrieve>.
- UN-GGIM (United Nations Committee of Experts on Global Geospatial Information Management). 2020. *Future Trends in geospatial information management: the five to ten year vision – Third Edition*. Último acceso: 10 de Noviembre, 2021, de https://ggim.un.org/meetings/GGIM-committee/10th-Session/documents/Future_Trends_Report_THIRD_EDITION_digital_accessible.pdf.
- UNHRC (United Nations Human Rights Council). 2020. *Detailed findings of the independent international factfinding mission on the Bolivarian Republic of Venezuela*. Agenda item 4. Human rights situations that require the Council's attention. Último acceso: 20 de Septiembre, 2020, de https://www.ohchr.org/Documents/HRBodies/HRCouncil/FFMV/A_HRC_45_CRP.11.pdf.
- UNODC (United Nations Office on Drugs and Crime). 2012. *Digest of Organized Crime Cases. A compilation of cases with commentaries and lessons learned*. Último acceso: 4 de Febrero, 2021, de https://www.unodc.org/documents/organized-crime/EnglishDigest_Final301012_30102012.pdf.
- UNODC (United Nations Office on Drugs and Crime). 2017a. *Estado Plurinacional de Bolivia. Monitoreo de Cultivos de Coca 2016*. La Paz. Último acceso: 20 de Agosto, 2020, de https://www.unodc.org/documents/crop-monitoring/Bolivia/2016_Bolivia_Informe_Monitoreo_Coca.pdf.
- UNODC (United Nations Office on Drugs and Crime). 2017b. *Colombia. Monitoreo de territorios afectados por cultivos ilícitos 2016*. Bogota. Último acceso: 20 de Agosto, 2020, de https://www.unodc.org/documents/colombia/2017/julio/CENSO_2017_WEB_baja.pdf.
- UNODC (United Nations Office on Drugs and Crime). 2017c. *Peru. Monitoreo de Cultivos de Coca 2016*. Lima. Último acceso: 20 de Agosto, 2020, de https://www.unodc.org/documents/crop-monitoring/Peru/Peru_Monitoreo_de_coca_2016_web.pdf.
- UNODC (United Nations Office on Drugs and Crime). 2018. *Peru. Monitoreo de Cultivos de Coca 2017*. Lima. Último acceso: 20 de Agosto, 2020, de https://www.unodc.org/documents/peruandecuador/Informes/monitoreo_coca/181213_InformeMonitoreo_2017Web.pdf.

- UNODC (United Nations Office on Drugs and Crime). 2019a. *Estado Plurinacional de Bolivia. Monitoreo de Cultivos de Coca 2018*. La Paz. Último acceso: 20 de Agosto, 2020, de https://www.unodc.org/documents/crop-monitoring/Bolivia/Bolivia_Informe_Monitoreo_Coca_2018_web.pdf.
- UNODC (United Nations Office on Drugs and Crime). 2019b. *Colombia. Monitoreo de territorios afectados por cultivos ilícitos 2018*. Bogota. Último acceso: 20 de Agosto, 2020, de <https://www.unodc.org/colombia/es/informe-de-monitoreo-de-territorios-afectados-por-cultivos-ilicitos-en-colombia-2018.html>.
- UNODC (United Nations Office on Drugs and Crime). 2019c. *World Drug Report 2019*. Último acceso: 19 de Mayo, 2020, de <https://wdr.unodc.org/wdr2019/>.
- UNODC (United Nations Office on Drugs and Crime). 2021a. *World Drug Report 2021*. Último acceso: 01 de Noviembre, 2021, de <https://wdr.unodc.org/>.
- UNODC (United Nations Office on Drugs and Crime). 2021b. *UNODC and illicit crop monitoring*. Último acceso: 10 de Febrero, 2021, de <https://www.unodc.org/unodc/en/crop-monitoring/index.html>.
- UNODC (United Nations Office on Drugs and Crime). 2021c. *Colombia. Monitoreo de territorios afectados por cultivos ilícitos 2020*. Bogota. Último acceso: 11 de Noviembre, 2021, de https://www.unodc.org/documents/crop-monitoring/Colombia/Colombia_Monitoreo_de_territorios_afectados_por_cultivos_ilicitos_2020.pdf.
- USGIF (United States Geospatial Intelligence Foundation). 2020. *Geospatial Intelligence & AI/ML Progress During a Pandemic*. Último acceso 05/11/2021, de https://s3-us-east-2.amazonaws.com/trjmag/wp-media-folder-trajectory-magazine/wp-content/uploads/2020/11/Geospatial_Intelligence_and_AI-ML_Progress_During_a_Pandemic.pdf.
- Vargas, R. 2020. *Automating Land Cover Change Detection: A Deep Learning Based Approach to Map Deforested Areas*. Tesis de Doctorado, (Instituto Nacional de Pesquisas Espaciais, São José dos Campos). <http://mtc-m21c.sid.inpe.br/col/sid.inpe.br/mtc-m21c/2020/06.09.11.59/doc/publicacao.pdf>.
- VoPham, T., Hart, J., Laden, F., & Chiang, Y.-Y. 2018. Emerging trends in geospatial artificial intelligence (geoAI): potential applications for environmental epidemiology. *Environ Health*, 17(1), 1-6. <https://doi.org/10.1186/s12940-018-0386-x>
- Warmerdam, F. 2008. The Geospatial Data Abstraction Library. *Open Source Approaches in Spatial Data Handling. Advances in Geographic Information Science* (págs. 87-104). Berlin: Springer. https://doi.org/10.1007/978-3-540-74831-1_5
- White, R. 2011. *Transnational Environmental Crime: Toward an eco-global criminology* (1st ed.). Tylor & Francis.
- Xia, G. S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J.,... & Zhang, L. 2018. DOTA: A Large-Scale Dataset for Object Detection in Aerial Images. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 3974-3983. <https://doi.org/10.1109/CVPR.2018.00418>
- Yang, Y., & Newsam, S. 2010. Bag-of-visual-words and spatial extensions for land-use classification. *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, 270-279. <https://doi.org/10.1145/1869790.1869829>
- Zabyelina, Y., & van Uhm, D. 2020. Illegal Mining: Organized Crime, Corruption and Ecocide in a Resource-Scarce World. <https://doi.org/10.1007/978-3-030-46327-4>
- Zhang, L., Zhang, L., & Du, B. 2016. Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, 4, 22-40. <https://doi.org/10.1109/MGRS.2016.2540798>
- Zhao, J., Zhang, Z., Yao, Y., Datcu, M., Xiong, H., & Yu, W. 2020. OpenSARUrban: A Sentinel-1 SAR Image Dataset for Urban Interpretation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 187-203. <https://doi.org/10.1109/JSTARS.2019.2954850>
- Zhu, X., Hu, J., Qiu, C., Shi, Y., Kang, J., Mou, L., Wang, Y. 2020. So2Sat LCZ42: A Benchmark Dataset for Global Local Climate Zones Classification. *IEEE Geoscience and Remote Sensing Magazine*. 1912.12171. <https://doi.org/10.1109/MGRS.2020.2964708>