

OPTIMIZACIÓN DE LA CLASIFICACIÓN SUPERVISADA DE IMÁGENES DE SATÉLITE
A PARTIR DE INFORMACIÓN DE RPAS PARA LA ACTUALIZACIÓN DE LA VARIABLE
COBERTURA DEL SUELO DEL MARCO MAESTRO RURAL AGROPECUARIO (MMRA)

LAURA ALEJANDRA MARTÍNEZ CONDE

20151032045

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
FACULTAD DEL MEDIO AMBIENTE Y RECURSOS NATURALES
INGENIERÍA TOPOGRÁFICA
BOGOTÁ D.C.
2020

OPTIMIZACIÓN DE LA CLASIFICACIÓN SUPERVISADA DE IMÁGENES DE SATÉLITE
A PARTIR DE INFORMACIÓN DE RPAS PARA LA ACTUALIZACIÓN DE LA VARIABLE
COBERTURA DEL SUELO DEL MARCO MAESTRO RURAL AGROPECUARIO (MMRA)

LAURA ALEJANDRA MARTÍNEZ CONDE

20151032045

Proyecto presentado como requisito para optar por el título de INGENIERO TOPOGRÁFICO

PROYECTO DE GRADO BAJO LA MODALIDAD DE PASANTÍA

Docente director: William Barragán Zaque

UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
FACULTAD DEL MEDIO AMBIENTE Y RECURSOS NATURALES INGENIERÍA
TOPOGRÁFICA BOGOTÁ D.C.

2020

CONTENIDO

1.	INTRODUCCIÓN.....	4
2.	OBJETIVOS.....	6
2.1	OBJETIVO GENERAL	6
2.2	OBJETIVOS ESPECÍFICOS	6
3.	MARCO TEÓRICO	7
3.1	Información estadística del sector agropecuario en Colombia.....	7
3.2	Teledetección.....	7
3.3	Clasificación de imágenes	11
3.4	Imágenes Sentinel.....	13
3.5	Sistema de aeronave pilotada a distancia – RPAS	15
4.	ACTIVIDADES Y PRODUCTOS GENERADOS	18
4.1	Recopilación de información del Estado del Arte.....	18
4.2	Desarrollo de la metodología y el algoritmo de clasificación	24
4.3	Clasificación de las imágenes mediante el script V03-03	32
5.	ANÁLISIS DE RESULTADOS.....	54
6.	EVALUACIÓN Y CUMPLIMIENTO DE LOS OBJETIVOS DE LA PASANTÍA.....	60
7.	CONCLUSIONES Y RECOMENDACIONES	61
8.	BIBLIOGRAFÍA.....	62

1. INTRODUCCIÓN

La gestión del capital natural de un país, en gran medida, depende del conocimiento del potencial de los recursos naturales existentes en el territorio que puedan proporcionar un flujo de bienes y servicios para satisfacer las necesidades de sus habitantes (F. Hernández & Martí, 2007). En Colombia, el Departamento Administrativo Nacional de Estadística - DANE, es la entidad que, en cumplimiento de una de sus funciones, se encarga de poner a disposición de la comunidad la información de los indicadores estadísticos referentes al sector agropecuario, incluyendo el área, la producción y el rendimiento de los cultivos (DANE, 2020)

Entre sus estadísticas en el sector agropecuario se encuentra la Encuesta Nacional Agropecuaria - ENA, que permite estimar estos atributos de los principales cultivos transitorios y permanentes, árboles frutales dispersos, áreas forestales y de pastos, en los treinta y dos departamentos del territorio colombiano (DANE, 2019a). Para llevar a cabo la ENA, se cuenta con un marco dual de áreas y listas que contiene información de las variables reportadas por el Tercer Censo Nacional Agropecuario realizado durante 2013 y 2014 (DANE, 2019b), dicho marco constituye lo que se conoce como Marco Maestro Rural y Agropecuario-MMRA que consolidó el DANE en el transcurso del 2015 al 2017 a partir de la información recogida por el mismo censo.

En cumplimiento del Plan de Acción de la FAO (2013) en el marco de la Estrategia Global para el Mejoramiento de las Estadísticas Agropecuarias y Rurales, el MMRA tiene por objeto constituirse como instrumento básico para el diseño muestral de las encuestas intercensales requeridas para proporcionar información más completa y precisa sobre el sector agropecuario y las zonas rurales que demande el país (DANE, 2016), es decir que, a partir de este, se faciliten los procesos de selección de la muestra para llevar a cabo la ENA, y que, durante los periodos intercensales, se puedan realizar comparaciones del cambio que han tenido las principales variables del sector.

Para cumplir con este objetivo y llevar a cabo una gestión de información estadística agropecuaria de confianza que permita la toma de decisiones tanto de los productores respecto a sus cultivos como de todo el país para el fortalecimiento de las políticas de desarrollo rural, seguridad alimentaria y competitividad del sector agropecuario (MADR, 2018), la Dirección de Geoestadística – DIG del DANE, encabezada por el grupo interno de trabajo Marco Geoestadístico Nacional, en una de sus funciones, se encarga de la actualización de la variable cobertura del suelo del MMRA; este proceso tiene como fin determinar el porcentaje existente de cada cobertura por unidad de observación (DANE, 2019b), es decir, la descripción cuantitativa de la variabilidad de los materiales físicos que se encuentran sobre la superficie de los conglomerados.

Las imágenes que se utilizan dentro del proceso de actualización son aquellas que están disponibles en la DIG, por ejemplo, imágenes PlanetScope que se han descargado y procesado, sin embargo, en algunas ocasiones, la fecha de captura de dicha imágenes no corresponde con el periodo del año en que se está actualizando el MMRA, lo que genera incertidumbre en la determinación de algunas coberturas que podrían haber cambiado, además, es importante resaltar que la resolución espacial de estas imágenes no es la apropiada para obtener una clasificación visual que responda a la necesidad real. Por otro lado, este proceso exige gran disposición de tiempo, lo que obliga a dar

prioridad a la actualización de los conglomerados de las investigaciones agropecuarias que realiza la entidad en el momento, y, como consecuencia, algunas unidades queden desactualizadas.

Por lo anterior, se hace necesario investigar el uso de tecnologías innovadoras y el desarrollo de metodologías que faciliten la obtención de los resultados y con un mayor grado de exactitud, de manera que la información estadística publicada responda a las necesidades de los distintos actores del sector agropecuario (CORPOICA, 2015), y, el auge del uso de los Sistemas de Aeronaves Pilotadas a Distancia, conocidos por las siglas RPAS, y otras imágenes de mejor resolución espacial junto con los avances en el desarrollo de los microsensores, han planteado el uso de estas plataformas como un nuevo concepto de teledetección donde se pueden superar las limitaciones de las imágenes satelitales de manera que se llegue a mejorar la clasificación final (Díaz García-Cervigón, 2015)

En el marco del convenio Colciencias-ENA, el GIT de Investigación y Desarrollo Estadístico del DANE, como respuesta a la necesidad de actualización del MMRA, propone una metodología de integración de imágenes de satélite con información de RPAS, que permite automatizar y dar mayor confiabilidad a la clasificación de las coberturas del suelo que conforman los conglomerados.

Para el desarrollo de la metodología de integración de la información, se hace uso de las herramientas que proporciona Google de manera gratuita, ya que se busca realizar estos procesos de una manera eficiente dado que el alcance que se aspira dentro de esta metodología es a nivel nacional. Estas herramientas corresponden en primera instancia, a un catálogo de imágenes satelitales obtenidas durante décadas a través de diferentes sensores, en segundo lugar, a los servidores de Google, que permiten realizar el procesamiento y análisis de la información satelital en la nube, y finalmente, a la plataforma *Google Earth Engine* – GEE, en la que se lleva a cabo el uso de distintos algoritmos de procesamiento de imágenes y de clasificación haciendo uso de dicho catálogo de imágenes satelitales, utilizando herramientas de procesamiento y análisis espacial con ayuda del lenguaje de programación JavaScript.

Para la clasificación, son empleadas las imágenes satelitales multiespectrales de media resolución espacial de la misión satelital Sentinel-2, e imágenes de alta resolución espacial del terreno capturadas con RPAS, las cuales son adquiridas a través de diferentes entidades como el IGAC e instituciones universitarias. Una vez obtenidas dichas imágenes, la información contenida en ellas fue utilizada para realizar su correspondiente integración con el mosaico satelital sentinel-2 y de esta forma mejorar sustancialmente la clasificación supervisada de imágenes en GEE y finalmente semiautomatizar el proceso de actualización de la variable cobertura del suelo del MMRA.

2. OBJETIVOS

2.1 OBJETIVO GENERAL

Apoyar al GIT de Investigación y Desarrollo Estadístico del Departamento Administrativo Nacional de Estadística en la elaboración de una propuesta experimental/exploratoria para automatizar la actualización de la variable cobertura del suelo del Marco Maestro Rural y Agropecuario basada en imágenes satelitales e información obtenida mediante Sistemas de Aeronaves Pilotadas a Distancia (RPAS).

2.2 OBJETIVOS ESPECÍFICOS

- Revisar y analizar fuentes que permitan establecer el uso de información de RPAS para integrarla con información satelital y mejorar la clasificación de coberturas.
- Contribuir al desarrollo de la metodología y el algoritmo de clasificación automatizada de coberturas del MMRA en la plataforma Google Earth Engine utilizando imágenes disponibles de Sentinel-2.
- Clasificar las coberturas de las áreas seleccionadas del MMRA a partir de las imágenes disponibles por medio de la plataforma Google Earth Engine.
- Contribuir a la generación de un artículo científico en el que se conceptualice la metodología propuesta y los resultados obtenidos.

3. MARCO TEÓRICO

3.1 Información estadística del sector agropecuario en Colombia

3.1.1 Encuesta Nacional Agropecuaria – ENA

La Encuesta Nacional Agropecuaria - ENA es una operación estadística que surge en la década de 1980 con el propósito de disponer de cifras estadísticas anuales del sector agropecuario. Es realizada en la actualidad por muestreo probabilístico a partir de un marco dual de áreas y listas, además de recolección de información mediante entrevista directa a los productores agropecuarios. Permite obtener estimaciones de variables estructurales y coyunturales de la actividad productiva agropecuaria del territorio colombiano y genera información relevante para la toma de decisiones por parte de los productores, las agremiaciones y las entidades del Estado. El marco dual de áreas y listas que permite llevar a cabo la ENA, se refiere a un marco que cuenta con información geográfica constituida por unidades de cobertura, unidades de observación y conglomerados, y con información alfanumérica constituida por atributos de identificación, ubicación y clasificación, de las variables reportadas por el Tercer Censo Nacional Agropecuario. Este marco constituye al Marco Maestro Rural y Agropecuario (DANE, 2020)

3.1.2 Marco Maestro Rural y Agropecuario - MMRA

El Marco Maestro Rural y Agropecuario – MMRA es un marco que surge en cumplimiento del primer objetivo propuesto por el Tercer Censo Nacional Agropecuario de “construir el marco estadístico agrícola, forestal, pecuario, acuícola, pesquero y de las actividades productivas no agropecuarias para las encuestas intercensales”. Tiene como base los elementos estructurales conocidos como marco de áreas y marco de listas, elementos que se interrelacionan de tal forma que las investigaciones pueden acceder a la información censal rural y agropecuaria que requieran de manera organizada y agrupada y georreferenciarla a través de las unidades de observación dispuestas para ello, ya sean áreas o listas. (DANE, 2019a)

Se encuentra delimitado por la totalidad del área rural del país que figura en los registros catastrales del Instituto Geográfico Agustín Codazzi - IGAC, para el año 2012, así como de los catastros descentralizados de Bogotá, Cali, Medellín y Antioquia, incluye, además, los Parques Nacionales Naturales que tienen área continental, permitiendo identificar el tipo de actividad que se desarrolla al interior de estas áreas protegidas. (DANE, 2019a)

3.2 Teledetección

La teledetección, es una técnica que permite adquirir información acerca de objetos presentes sobre la superficie terrestre a través de la captura de imágenes de la tierra mediante sensores instalados a bordo de plataformas satélites o plataformas aéreas para su posterior procesamiento e interpretación; para ello se requiere, primero, disponer de una fuente que ilumine y suministre de energía al objeto de interés, ya sean bosques, ciudades, mares, cultivos, u otros, dependiendo del sensor, dicha fuente puede ser el sol o una fuente de radiación artificial. (García Labrador et al., 2012).

El sensor que capta la radiación electromagnética emitida por fuentes naturales se denomina captador pasivo, y aquel que posee fuentes internas que generan artificialmente la radiación se denomina captador activo. (Sarría, 2006). Esta radiación, en su recorrido hacia la tierra, atraviesa e interacciona con la atmósfera para luego alcanzar la superficie terrestre, los objetos que en ella se encuentren, puede reflejar, absorber o transmitir esta radiación de forma diferente de acuerdo con sus características físicas y químicas, diferencias que posteriormente permiten distinguir un objeto de otro (Sarría, 2006).

El sensor recoge y graba esa radiación reflejada tanto por los objetos presentes en la superficie terrestre como por la atmósfera en formato digital, la transmite a una estación de recepción y procesamiento, y los datos se convierten en imágenes digitales (Cede et al., 2011).

La historia de esta técnica da inicio en el año 1946, cuando fueron lanzados desde Nuevo México los primeros cohetes V-2 con pequeñas cámaras fotográficas instaladas en ellos como sensores remotos (Pérez Gutiérrez & Muñoz Nieto, 2006), y, partir de ese momento, han sido lanzadas diferentes misiones a bordo de otros cohetes, misiles balísticos y satélites con el fin de realizar una observación sistemática de la Tierra. El producto generado a partir de los sensores, es una imagen digital tipo raster donde cada píxel tiene asignado uno o varios valores numéricos, denominado Valores Digitales, que hacen referencia a la energía media recibida dentro de una determinada banda espectral (Yébenes Gómez & Giner Sotos, 2010). La calidad de las imágenes obtenidas con estos instrumentos ha ido mejorando con el tiempo de forma exponencial, aumentando con ello el interés de la comunidad científica y la sociedad general por el uso de esta información (García Labrador et al., 2012).

3.2.1 Resolución de los sensores remotos

La radiación, tanto emitida como reflejada, es un fenómeno continuo en cuatro dimensiones, espacio, tiempo, longitud de onda, y radiancia; el nivel de detalle con que un sensor captura esta información se denomina resolución, y, por lo tanto, se define un tipo de resolución para cada dimensión: resolución espacial, resolución temporal, resolución espectral y resolución radiométrica (Bense, 2007).

La resolución espacial es aquella que designa el tamaño del objeto más pequeño que puede ser distinguido en una imagen, por lo tanto, se refiere al tamaño de cada uno de los pixeles que la conforman, cuanto menor es su tamaño, mayor es el valor de la resolución (Paz Cerda, 2004). La forma más común de determinarla es mediante el IFOV - *Instant Field of View*, el cual define el área de terreno captado y es predispuesto por las características geométricas del sistema sensor (INGEOMINAS, 2008).

Como consecuencia de que se afirma que un objeto solo es reconocible en una imagen si este tiene un tamaño mayor al del píxel, la resolución espacial está relacionada de manera directa con la escala de trabajo y con la fiabilidad obtenible de una interpretación. Entre más pequeñas sean las dimensiones del píxel, la probabilidad de que éste sea compuesto por más de una cubierta también será menor (INGEOMINAS, 2008). Un píxel mixto, viene definido por una señal que es mezcla de las señales de los diversos elementos que le componen. (INGEOMINAS, 2008).

La resolución temporal, indica el intervalo de paso del satélite sobre la vertical de un punto, por lo tanto, es una medida que determina la frecuencia con la que un satélite adquiere imágenes de una misma zona (Tristan et al., 2008). El ciclo de cobertura está condicionado por el tipo de órbita, el

ángulo de barrido y la velocidad del satélite. Altas resoluciones temporales son importantes en el monitoreo de eventos que cambian en períodos relativamente cortos, como inundaciones, incendios, calidad del agua en el caso de contaminaciones, desarrollo de cosechas, entre otros (Bense, 2007). Asimismo, en áreas con cubiertas nubosas casi constantes, como por ejemplo las selvas tropicales, períodos cortos de visita, es decir altas resoluciones temporales, aumentan la probabilidad de obtener imágenes satisfactorias. (UNCPBA, 2012)

La resolución espectral, indica el número y ancho de las bandas en las que el sensor puede captar la radiación electromagnética. Mayor número de bandas implica mayor exactitud de los datos, ya que cada banda constituye una variable para caracterizar la superficie captada (Díaz García-Cervigón, 2015). Dependiendo de la aplicación, pueden seleccionarse sensores con bandas relativamente estrechas o anchas, generalmente, cuando se tienen muchas bandas, estas corresponden con intervalos de ancho menor, y dan al sensor una resolución espectral mayor. (TELEDET, 2020)

La resolución radiométrica, es la capacidad del sensor para detectar variaciones en la radiancia espectral que recibe (León, 2002). Consiste en el número de niveles digitales utilizados para expresar los datos recogidos por el sensor, representados en tonos de gris. En general, cuando mayor es el número de niveles, mayor es el detalle con que se podrá expresar dicha información (Bense, 2007). Al igual que en los otros tipos de resolución, cuanto mayor sea la precisión radiométrica, mejor podrá interpretarse la imagen. El número de niveles más adecuado depende del método que se siga en la interpretación, si es visual, conviene considerar que el ojo humano difícilmente percibe más de 64 niveles de gris, y no más de 200.000 tonalidades (INGEOMINAS, 2008). Esta resolución suele expresarse mediante el número de bits necesarios que se precisan para almacenar cada píxel, definidos como unidad fundamental de un sistema binario, pudiendo poseer sólo los valores 1 y 0. (Sarría, 2006)

3.2.2 Interacción de la vegetación con la radiación

La vegetación refleja la radiación de una forma u otra dependiendo de varios factores: sus características morfológicas, como la altura, perfil, grado de cobertura del suelo; su estado fenológico, forma y contenido de humedad; e incluso la ubicación geográfica, pendiente, orientación, asociación con otras especies o geometría de la plantación (Martínez Barbáchano & Solís Miranda, 2018).

De acuerdo con Hernández & Montaner, (2009) la vegetación sana tiene una reflectividad baja en el espectro visible, aunque con un pico en la banda verde debido a la clorofila. La reflectividad es muy alta en el infrarrojo cercano debido a la escasa absorción de energía por parte de las plantas en esta banda. En el infrarrojo medio, hay una disminución especialmente importante en aquellas longitudes de onda en las que el agua en los tejidos de la planta absorbe la energía (I.B. et al., 1999). En el caso de la vegetación enferma, como las hojas tienden a perder clorofila, la reflectividad aumenta en las bandas roja y azul; por el contrario, en el infrarrojo cercano, se produce una reducción de la reflectividad como consecuencia de un deterioro en la estructura celular de la hoja. (Gallo, 2002)

Asimismo, Sarría, (2006) expone que la reflectividad de una planta depende de su contenido en agua, cuando el contenido de esta aumenta, la reflectividad disminuye como consecuencia de que incrementa la absorción de la radiación por parte del agua que contiene la planta.

Reuter, (2009) Explica que la cantidad de energía que viaja con longitudes de onda mayores, como en el infrarrojo cercano, es demasiado baja para ser aprovechada por la vegetación, por lo que sólo se absorbe en torno al 5 %, el 40% se transmite y el 55% se refleja, sin embargo, la energía transmitida acabará llegando a otra hoja situada por debajo, por lo tanto, el 55% de ese 40% transmitido se volverá a reflejar.

Por otro lado, conforme a INGEOMINAS, (2008), en la teledetección las plantas se presentan como hojas de diferentes formas y tamaños, con espacios entre las mismas, además de muchas otras variaciones y, todas las interacciones individuales en estas estructuras complejas, pueden interferir y posteriormente ampliar el rango de las respuestas, porque el sensor no mide hojas aisladas, sino agrupaciones de las mismas que constituyen masas de vegetación. Además, intervienen otros elementos adicionales que complican su detección, tales como el ángulo de elevación solar, directamente relacionado con la proporción de sombras que detecta el sensor, las características del suelo que sustenta la vegetación, la geometría de la cubierta vegetal o el estado fenológico en que se encuentra la vegetación. (Reuter, 2009)

3.2.3 Índices de vegetación

Con el fin de realzar la vegetación existente en una imagen en función de su respuesta espectral y atenuar los detalles de otros elementos presentes, es posible realizar operaciones algebraicas entre las bandas espectrales que la componen y, así, generar una nueva imagen con las características deseadas donde se destaquen gráficamente determinados píxeles relacionados con parámetros de las coberturas vegetales (Sancha Navarro, 2010). Dicho conjunto de operaciones, que pretende extraer de los valores de reflectividad la información relacionada con la vegetación, se denomina como índice de vegetación (Muñoz Aguayo, 2013).

Existe una gran variedad de índices de vegetación que tienen como objetivo aportar información relevante acerca de la misma (C. Hernández et al., 2014); gran parte de estos índices tienen su sustento, en mayor medida, en la interacción entre la vegetación y la energía electromagnética de las bandas del espectro rojo e infrarrojo cercano, debido a que las plantas absorben muy poca energía en estas bandas (Paz Pellat et al., 2014).

A modo de objetivo de estos índices, cabe mencionar la posibilidad que ofrece de integración en una única imagen de valores altamente correlacionados con parámetros de verdor o vigor, índice de área foliar -LAI, productividad, concentración de clorofila, eficiencia fotosintética y contenido de agua entre otros. (Paruelo, 2008)

A la hora de trabajar o realizar cálculos de los diferentes índices espectrales, es necesario trabajar con valores de reflectancia en superficie. Para ello, es necesario efectuar la transformación inicial de Niveles Digitales a radiancia. Posteriormente, los valores de radiancia se transforman en valores de reflectividad en TOA - *Top of Atmosphere*, y tras realizar la corrección atmosférica y topográfica pertinente se obtienen los valores finales de reflectancia en superficie (Caseras Saez, 2017).

3.3 Clasificación de imágenes

A partir de las imágenes obtenidas mediante sensores remotos, es posible categorizar cada uno de los píxeles en una o diversas clases temáticas y, posteriormente, generar mapas temáticos (Cabrera et al., 2011). Los programas de clasificación llevan a cabo este proceso mediante aprendizaje autónomo, por lo tanto, se limitan a agrupar conjuntos de objetos en clases a través del uso de técnicas de inteligencia artificial o estadísticas (Arango et al., 2005).

El proceso de clasificación involucra dos fases, la definición de las clases y la adjudicación de cada uno de los píxeles de la imagen a dichas clases. En función de la forma en que se obtienen las estadísticas de las clases, los métodos de clasificación se dividen en dos tipos, métodos supervisados y métodos no supervisados (Yébenes Gómez & Giner Sotos, 2010).

- Clasificación no supervisada:

Es un clasificador que usa algoritmos para agregar píxeles en grupos naturales o clúster con características espectrales similares presentes en la imagen (Ingeoexpert, 2020). En este caso, la influencia que puede ejercer el operador no va más allá de seleccionar el algoritmo de clasificación a utilizar y la comprobación de la correspondencia entre la clase espectral y la clase temática, debiendo llevar a cabo varias veces el mismo procedimiento hasta obtener un resultado final considerado como válido (Monterroso, 2013). En esta clasificación no se establece previamente ninguna clase, aunque sí es necesario determinar el número de ellas que se desea establecer. El resultado de esta clasificación no supervisada muchas veces es utilizado para definir las áreas de entrenamiento en una posterior clasificación supervisada. (Arango et al., 2005)

- Clasificación supervisada:

Para realizar esta clasificación el analista debe tener un conocimiento previo del área de estudio. Involucra las etapas de entrenamiento, clasificación y operaciones de post clasificación.

En la etapa de entrenamiento, el analista selecciona áreas representativas de identidad conocida sobre la zona de interés mediante rectángulos o polígonos, generando una descripción numérica de los atributos espectrales de cada categoría de cobertura de la tierra (Bense, 2007); estos datos constituyen los “datos de entrenamiento”. Para llevar a cabo la etapa de clasificación, el algoritmo seleccionado adjudica cada uno de los píxeles de identidad desconocida a alguna clase de acuerdo a su mayor semejanza con los datos de entrenamiento (Sarría, 2006). Las operaciones de post clasificación involucran el análisis de precisión y la verificación de los resultados; se realizan con el fin de estimar la exactitud de la clasificación al definir el grado de concordancia entre las clases asignadas por el clasificador y los datos definidos como verdaderos (Ordoñez & Serna, 2015). El método más apropiado para la evaluación es partir de matrices de confusión, ya que comparan los resultados de la clasificación con los datos de validación, y refleja además el desempeño del algoritmo empleado. (Yébenes Gómez & Giner Sotos, 2010)

En casos en que la definición de las áreas de entrenamiento y la correspondiente codificación de las clases de interés sea satisfactoriamente realizable, el método supervisado es superior al no supervisado, cabiendo mencionar que este posee un enfoque subjetivo ya que el analista intenta

definir clases temáticas compuestas, a veces, por varias clases espectrales (Monterroso et al., 2018); en el caso de áreas de terreno complejas, donde el analista tenga dificultad para definir las áreas de entrenamiento dada la variabilidad de la respuesta espectral dentro de cada clase, el método no supervisado resulta superior al supervisado, además de tener el potencial para revelar clases temáticas no previstas previamente. (TELEDET, 2020)

3.3.1 Clasificación basada en píxeles

La forma tradicional de extracción de información temática en imágenes satelitales es la clasificación basada en píxeles, que se ha empleado de forma generalizada en todo tipo de proyectos desde los inicios de la interpretación de imágenes asistida por computador, en la década de 1960 (Chuvieco, 2008). Para ello, el patrón espectral presente dentro de la información de cada píxel se usa como una base numérica para la categorización, lo que ocasiona que diferentes tipos de objetos presenten diferentes combinaciones de Niveles Digitales basados en su inherente reflectancia espectral y propiedades de emisividad (MAE-MAGAP, 2015). Existen diferentes tipos de patrones que permiten la categorización de los píxeles de una imagen, como ser patrones espectrales, temporales y espaciales (Ingeoexpert, 2020).

Para generar información temática, los píxeles individuales que conforman la imagen deben ser discretizados o agrupados en categorías de interés que son definidas por el intérprete de acuerdo a sus objetivos y necesidades (Ardila et al., 2005). Este agrupamiento de píxeles se realiza tradicionalmente mediante un proceso de agrupamiento de datos estadístico (*clustering*) en el que cada píxel es asignado a un grupo o categoría de acuerdo a su correspondiente nivel digital, es decir se realiza una clasificación basada en respuesta espectral de píxeles individuales (Arango et al., 2005).

3.3.2 Algoritmos de clasificación - Machine Learning

Machine Learning es un campo de la inteligencia artificial que comprende un conjunto de técnicas para que, a través del entrenamiento con grandes volúmenes de datos, se creen sistemas de aprendizaje automático; dichos sistemas utilizan algoritmos programados que reciben y analizan los datos y a partir de ellos predicen comportamientos futuros. (Russo et al., 2016). Estos sistemas de aprendizaje automático se clasifican en tres grupos:

Primero, el aprendizaje supervisado, en el que la máquina aprende a través de datos de ejemplo y repuestas de destino asociadas, para luego predecir la respuesta correcta cuando se presenta con nuevos ejemplos; el algoritmo realiza predicciones y es corregido por el operador hasta que el algoritmo alcanza un alto nivel de precisión y rendimiento (González, 2018). Segundo, el aprendizaje no supervisado, en el que la máquina aprende de ejemplos simples sin ninguna respuesta asociada, dejando que el algoritmo determine los patrones de datos por sí mismo, a medida que evalúa más datos, su capacidad para tomar decisiones sobre los mismos mejora gradualmente y se vuelve más refinada (Oracle, 2020). Tercero, el aprendizaje reforzado, en el que no hay una clave de respuesta, y, en ausencia de datos de capacitación, el algoritmo aprende de la experiencia, recoge los ejemplos de los procesos de aprendizajes reglamentados, en los que se proporcionan algoritmos de aprendizaje automáticos con un conjunto de acciones, parámetros y valores finales. (Redacción APD, 2019). Derivado de la revisión de los distintos métodos de *machine learning* aplicados a la clasificación supervisada basada en píxeles de imágenes satelitales fue evidenciado que el que posee mejor comportamiento es *Random Forest* (Belgiu & Csillik, 2017;

Chen et al., 2017; Zurqani et al., 2017)., siendo este el seleccionado en el flujo metodológico del documento presentado.

3.3.2.1 Algoritmo de clasificación de aprendizaje supervisado - *Random Forest*

Random Forest es un algoritmo que puede ser utilizado tanto para tareas de regresión como de clasificación, en el que se ejecutan varios algoritmos de árboles de decisión en lugar de uno solo con el objetivo de clasificar un nuevo objeto basado en atributos; para ello, cada árbol da una clasificación y finalmente la decisión con mayor número de votos es la predicción del algoritmo. (González, 2018)

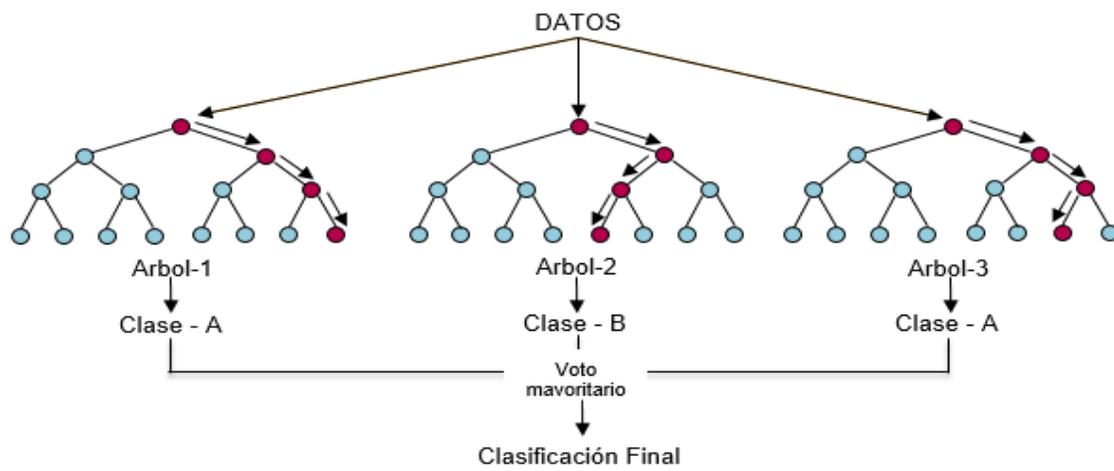


Ilustración 1. *Random Forest*. Fuente: Adaptado de (Koehrsen, 2017)

El clasificador *Random Forest* usa un conjunto de clasificadores (coberturas definidas) para generar la regresión en el árbol de decisión como se aprecia en la Ilustración 1, es muy sensible a información de entrenamiento no adecuada, pero robusto en situaciones de sobreajuste (Belgiu & Drăgu, 2016). Es decir, en gran medida la exactitud temática del clasificador *Random Forest* depende de la experiencia del intérprete y la generación adecuada de los polígonos o áreas de entrenamiento.

3.4 Imágenes Sentinel

Copernicus, es el programa de observación de la Tierra diseñado para proporcionar información precisa, actualizada y de fácil acceso para mejorar la gestión del medio ambiente, comprender y mitigar los efectos del cambio climático y garantizar la seguridad ciudadana. Esta iniciativa está liderada por la Comisión Europea - EC en colaboración con la Agencia Espacial Europea – ESA (Albacete, 2019).

La ESA ha desarrollado seis misiones Sentinel para atender a las necesidades del programa Copernicus, cada misión se basa en una constelación de dos satélites para cumplir con los requisitos de revisión y cobertura (ESA, 2020):

- Sentinel-1, es una misión de imágenes de radar de órbita polar, capaz de observar la superficie de la Tierra en cualquier condición meteorológica, de día o de noche.

- Sentinel-2, es una misión de imágenes de alta resolución multiespectral de órbita polar para monitorizar la superficie de la Tierra.
- Sentinel-3, transporta instrumentos para llevar a cabo una larga serie de estudios de la Tierra y de sus océanos.
- Sentinel-4 y Sentinel-5, miden la composición de la atmósfera desde la órbita geostacionaria y una órbita polar, respectivamente.
- Sentinel-6, lleva un altímetro de radar para medir la altura global de la superficie del mar, principalmente para la oceanografía operacional y para estudios climáticos.

3.3.2 Imágenes Sentinel-2

La misión Sentinel-2 se basa en una constelación de dos satélites idénticos en la misma órbita, separados 180° para garantizar una cobertura y entrega de datos óptimas. Juntos cubren todas las superficies terrestres de la Tierra, grandes islas, aguas continentales y costeras cada cinco días en el ecuador, en condiciones sin nubes, y de 2 a 3 días en latitudes medias. Sentinel-2A se lanzó el 23 de junio de 2015 y Sentinel-2B lo siguió el 7 de marzo de 2017. Una vez lanzado, la Comisión Europea asume la propiedad de los satélites (ESA, 2015).

Proporciona principalmente información para prácticas agrícolas y forestales y para ayudar a gestionar la seguridad alimentaria (Zhao et al., 2019), ya que estas imágenes satelitales se pueden usar para estimar índices como el LAI o características de las plantas como los contenidos de clorofila a nivel de cubierta, a nivel de hoja, y a nivel del área foliar (Pasqualotto et al., 2019).

Las imágenes generadas poseen 13 bandas espectrales que garantizan series temporales consistentes, muestran la variabilidad en las condiciones de la superficie terrestre y minimizan los efectos producidos por la variabilidad atmosférica. (ESA, 2020)

Tabla 1. Bandas de Sentinel-2. Fuente: Adaptado de (ESA, 2015)

Número de banda	Nombre	Longitud de onda central (µm)	Banda ancha (µm)	Resolución espacial (m)
1	Coastal aerosol	0.443	0.27	60
2	Blue	0.49	0.98	10
3	Green	0.56	0.45	10
4	Red	0.665	0.38	10
5	Vegetation Red Edge	0.704	0.19	20
6	Vegetation Red Edge	0.74	0.18	20
7	Vegetation Red Edge	0.783	28	20
8	NIR	0.842	0.145	10
8a	Vegetation Red Edge	0.865	0.33	20
9	Water Vapour	0.945	0.26	60
10	SWIR -Cirrus	1.375	0.75	60
11	SWIR	1.61	0.143	20
12	SWIR	2.19	0.242	20

3.5 Sistema de aeronave pilotada a distancia – RPAS

En Colombia, la Aeronáutica Civil (2018) define a un RPAS como una aeronave no tripulada que es pilotada desde una estación de pilotaje a distancia, tiene un “piloto remoto”, emplazado en una “estación de piloto remoto” ubicada fuera de la aeronave, es decir en tierra, en barco, en otra aeronave, en el espacio, etc., quien monitorea la aeronave en todo momento y tiene responsabilidad directa en la conducción segura de la aeronave durante todo su vuelo. Un RPAS puede poseer varios tipos de tecnología de piloto automático, pero, en todo momento, el piloto remoto puede intervenir en la gestión del vuelo.

Tabla 2 Diferentes definiciones de Drones. Fuente: Elaboración Propia

ACRÓNIMO	TÉRMINO	DEFINICIÓN
UAV	Unmanned Aerial Vehicle	Vehículo aéreo no tripulado
UAS	Unmanned Aerial System	Sistema aéreo no tripulado
RPA	Remotely Piloted Aircraft	Aeronaves controlados de forma remota
RPAS	Remotely Piloted Aircraft System	Sistema aéreo tripulado de forma remota

Los RPAS tienen definidos dos segmentos:

- Segmento de vuelo: formado por el vehículo aéreo y los sistemas de recuperación (aterrizaje sobre ruedas o patines, red, cable, paracaídas, etc.).
- Segmento de Tierra: formado por la estación de control (está en tierra, envía y recibe la información enviada por los drones) y el sistema lanzador (pueden ser hidráulicos, neumáticos, etc.).

En la obtención de datos con un RPAS es necesario que existan los siguientes sistemas (Díaz García-Cervigón, 2015):

- Plataforma de vuelo: el propio vehículo y su carga, es decir, la cámara fotográfica y otros sensores.
- Sistema de control de vuelo: formado por receptores GPS y EGNOS, integrados en la plataforma de vuelo. Georreferencia la información obtenida y controla el seguimiento del vehículo aéreo.
- Sistema de lanzamiento y recuperación: controla el despegue y el aterrizaje del vehículo aéreo.
- Sistema de comunicaciones: comunica la estación de control con el RPAS, transmite la información adquirida durante el vuelo.

Los RPAS se pueden clasificar en función de su tipo de sustentación: de ala fija y de ala rotatoria. Las plataformas de vuelo de ala rotatoria (multirrotores) son vehículos que tienen la capacidad de despegar y aterrizar en vertical desde el propio suelo sin necesitar una pista específica para ello (Oñate de Mora, 2015). Son autónomos, o sea que pueden volar con ayuda de los sistemas de control y navegación y únicamente necesitan la ayuda del piloto para despegar y aterrizar, pueden permanecer parados en el aire para tomar las imágenes y se pueden acercar al objeto a fotografiar con mucha precisión.



Ilustración 2. RPAS ala rotatoria Phantom 3. Fuente: (EMBENTION, 2016)

Las plataformas de vuelo de ala fija (tipo avión) tienen la ventaja de ser capaces de volar a altas velocidades durante largo tiempo. Para ello cuentan con elementos aerodinámicos más simples. La mayoría despegan a mano o con un lanzador y no requieren pista para el aterrizaje. Tienen menor capacidad de carga y no se pueden detener en el aire para obtener las imágenes, sin embargo, tienen mayor autonomía y menor riesgo de accidente. (Robles et al., 2016)



Ilustración 3. RPAS Ala fija NM& F300. Fuente: (EMBENTION, 2016)

La Resolución 04201 del 27 de diciembre del 2018 de la Aeronáutica Civil, establece parámetros e imparte instrucciones de cumplimiento en cuanto a los requisitos de aeronavegabilidad y operaciones necesarias para obtener permiso para la realización de operaciones con Sistemas de Aeronaves Pilotadas a Distancia - RPAS diferentes a las de recreación y deporte, clasificándolos de acuerdo con el riesgo operacional, así:

- Clase A (abierta). Corresponde a la operación de UAS¹ que se encuentren dentro de las limitaciones establecidas en el párrafo (a) de la sección 2.2. del apéndice 13 del RAC 91 y que cuenten con un MTOW superior a 250 gr y de hasta 25 kg, por lo cual no requieren de autorización de la UAEAC, dado que su operación representa un mínimo riesgo.
- Clase B (reculada). Corresponde a la operación de UAS que no se encuentren dentro de la Clase A por desviación de las limitaciones establecidas en el párrafo (a) de la sección 2.2. del apéndice 13 del RAC 91 o que cuenten con un MTOW superior a 25 kg y de hasta 150 kg, por lo cual requerirán siempre de autorización de la UAEAC, aun cuando su operación podría implicar un muy bajo riesgo.
- Clase C (certificada- RPAS). Corresponde a la operación de UAS por fuera de las limitaciones establecidas en la sección 3.2. del apéndice 13 del RAC 91, con MTOW superior a 150 kg, sobrevuelos internacionales y aquellos que prestaren servicios de

¹ La expresión UAS incluye las expresiones UAV, UAS, RPA, RPAS, VANT, DRON o DRONE, indiferentemente de su principio de vuelo o propulsión.

transporte, para los cuales, por ahora, no se autoriza su operación en el espacio aéreo donde tenga jurisdicción el Estado colombiano, aun cuando su operación pudiera ser de bajo riesgo.

Según las disposiciones del RAC - Reglamento Aeronáutico Colombiano, los pilotos de RPA, deben recibir un curso de formación para piloto privado en un centro de instrucción aeronáutico aprobado por la UAEAC, ya sea en una universidad nacional o extranjera, que posea una facultad de ingeniería aeronáutica o aeroespacial y un programa de educación continuada para RPAS. Los principios básicos de dichos programas deben ser las normas generales de operación de aeronaves RPA, la clasificación de espacios aéreos y servicios de tránsito aéreo que en ellos se presta, las disposiciones actuales sobre RPAS, los principios básicos de aerodinámica y vuelo, la meteorología aeronáutica, las comunicaciones aeronáuticas, el conocimiento de la aeronave por operar, los sistemas de gestión de seguridad operacional (SMS), la clasificación de los RPAS, el concepto de aeronavegabilidad. (Ávila Herrera, 2017)

4. ACTIVIDADES Y PRODUCTOS GENERADOS

4.1 Recopilación de información del Estado del Arte

Se llevó a cabo una extensa revisión de la literatura existente para los siguientes temas:

- Integración de imágenes satelitales con imágenes capturadas con RPAS
- Uso de la plataforma Google Earth Engine para la clasificación de imágenes
- Uso de algoritmos de Machine Learning para la clasificación supervisada

Lo anterior con el fin de construir el Estado del Arte para el documento metodológico y el artículo científico, además de contar con información de referencia que permitiera permitan establecer el uso de la información y los procesos a ser llevados a cabo para la efectiva integración.

Se analizaron y documentaron un total de ciento treinta y nueve (139) documentos considerados como relevantes para la investigación, entre ellos, artículos, libros, tesis, páginas web, y demás, de los cuales fueron citados ciento doce (112) en el documento metodológico y cuarenta y nueve (49) en el artículo científico.

4.1.1 Resultados

Estado del Arte

La teledetección es una herramienta que, en una de sus aplicaciones, sirve como base de la toma de decisiones para la gestión eficiente de los productos agrícolas; las imágenes procesadas se interpretan de forma visual o digital para extraer información de las coberturas presentes, y, a partir de ello, adquirir un mayor conocimiento de los recursos bajo estudio y aplicarlo en la resolución de problemas. (García Labrador et al., 2012)

La teledetección ha sido utilizada en la mayoría del mundo, lo cual se ve reflejado en la gran cantidad de imágenes que se generan desde diferentes satélites construidos y puestos en órbita por países como Estados Unidos, Israel, Corea o India (García Labrador et al., 2012), por nombrar algunos, y que se utilizan con el mismo objetivo, determinar las coberturas del suelo, sus características y monitorear sus cambios (FAO, 1990).

Además, los trabajos actuales en el campo de la teledetección se están centrando en poder relacionar la información fisiológica de los cultivos con la energía reflejada por los mismos a diferentes longitudes de onda, con el fin de diferenciar al cultivo propiamente dicho, la maleza, determinar los niveles de nutrientes, o incluso la detección de enfermedades y plagas. (González & Cancela, 2018) A partir de estos trabajos, se han demostrado las ventajas que tiene hacer uso de datos satelitales para la clasificación digital de coberturas, entre ellas, que la información multiespectral contenida en la imagen permite la extracción de información adicional que el ojo humano no captura con sus receptores rojo, verde y azul, esto gracias a que los sensores que captan la información reflejada por la superficie terrestre lo hacen en diferentes bandas del espectro electromagnético (García Labrador

et al., 2012), además, el satélite cuenta con la gran ventaja de no tener que desplazarse al campo en cuestión y permite abarcar grandes extensiones de terreno, y la frecuencia de paso del satélite que casi siempre se repite, permite obtener una colección de imágenes multitemporales que posibilitan observar los cultivos en sus diferentes etapas (Riaño M, 2002).

Por ejemplo, Willington et al. (2013) determinaron que, al clasificar usos del suelo agrícola con imágenes Landsat en zonas heterogéneas de la región central de Córdoba, Argentina, considerando aspectos como los distintos cultivos, fechas de siembra y estado fenológico, la aplicación de imágenes satelitales multitemporales es el procedimiento más adecuado para mejorar la discriminación de las coberturas, concluyendo que el análisis multitemporal permite una precisión de la clasificación más eficiente respecto de una clasificación realizada con datos una sola escena, logrando valores catalogados como excelentes en todas las clases de cobertura cuando se quiere identificar soja y maíz.

Sin embargo, Berni et al. (2009) concluyen que esta imagen general del terreno está condicionada por las limitantes en cuanto a resolución espacial y temporal de las imágenes disponibles se refiere, ya que consideran que estas resoluciones son inadecuadas al permitir obtener un resultado de carácter más general, inapropiado para parcelas pequeñas; de igual manera, León, (2002) señala las desventajas de las imágenes satelitales y como estas han perjudicado en gran medida el uso de dichos insumos para el manejo de cultivos, resaltando las frecuencias de revisión insuficientes, la necesidad de correcciones geométricas y radiométricas, la confusión que generan diferentes fenómenos presentes en la imagen con las mismas características espectrales y la necesidad de calibrar la imagen si se trata de áreas desconocidas por el intérprete, concluyendo que estas imágenes generalmente no son apropiadas para mapas detallados; por lo tanto, a pesar de que los métodos convencionales de clasificación de coberturas se realizan por medio de imágenes satelitales, este procedimiento puede resultar dispendioso e impreciso para aplicaciones en la agricultura (Gevaert et al., 2015).

Además, otra consideración importante es que la información espectral contenida en los píxeles de vegetación normalmente no es suficiente como para identificar especies o tipos de cubiertas (Sebem et al., 2005), y, dependiendo de la resolución espacial y del FOV o campo de visión del satélite, se determina si el pixel contiene únicamente información del cultivo o incluye una mezcla radiométrica de sus vecinos (Morales Cruz, 2013), por lo tanto, como pocas zonas tienen homogeneidad total, lo que se ocasiona es que la extracción de información útil de las imágenes de satélite se convierta en el principal problema al realizar teledetección, es decir, si se trata de cultivos no homogéneos, este método no funciona correctamente.

El auge del uso de los sistemas aéreos tripulados de forma remota, conocidos por las siglas RPAS, junto con los avances en el desarrollo de los microsensores, han planteado el uso de estas plataformas como un nuevo concepto de teledetección donde se pueden superar las limitaciones en distintos niveles de resolución. (Díaz García-Cervigón, 2015). Su potencialidad en la agricultura para captar información viene dada por la precisión espacial con la que se puede tomar el dato y por la disponibilidad temporal de ese dato (Montesinos Aranda, 2015).

Los RPAS, generalmente conocidos como drones, son vehículos aéreos no tripulados controlados a distancia capaces de portar distintos sensores de medición. De manera general, existen dos tipos de plataforma de RPAS, de ala fija y de ala rotatoria (Oñate de Mora, 2015), en función del objetivo

buscado, será necesario el empleo de una plataforma u otra. Los condicionantes principales para seleccionarla serán la superficie para estudiar, la resolución espacial requerida y las condiciones para el despegue y el aterrizaje en la zona (Díaz García-Cervigón, 2015)

Algunos estudios han comparado los RPAS con otros sistemas de teledetección y con otros vehículos aéreos, encontrando que proporcionan ciertas ventajas al momento de capturar y analizar la información contenida en las imágenes para cuantificar y caracterizar las áreas agrícolas de una región (Barreiro & Valero, 2014), como, por ejemplo, la posibilidad de determinar la resolución espacial al variar la altura de vuelo o la capacidad superior de adquisición de las imágenes al poderse capturar incluso en días nublados; esto da como resultado un proceso de recolección de datos que permite la adquisición de información precisa y frecuente del estado de los cultivos que requiere un país y sus productores (Montesinos Aranda, 2015).

Berni et al., (2009) en su estudio, consideran que los RPAS cubren perfectamente ciertas problemáticas que tienen las plataformas satelitales tales como la monitorización continua o en tiempo real, resolución espacial y la considerable mejora para este tipo de estudios como es la capacidad de revisita o resolución temporal.

De acuerdo con Méndez et al., (2020) muchas de las empresas del sector agropecuario adquieren drones para dar resultados agronómicos en lo que puede referirse a conocer mejor la variabilidad de los lotes, realizar un seguimiento de los cultivos de mejor manera y con mayor posibilidades de encontrar problemas rápidamente como puede ser un ataque de plagas y enfermedades, sectorizar ambientes de malezas, detectar fallas de siembra o fertilización si las hubiese, o para hacer un simple recorrido al campo desde la altura en momentos donde es difícil desplazarse al interior de los cultivos.

Barreiro & Valero (2014) consideran que, a pesar de que se tiene conocimiento de que las variables biológicas y físicas de un campo generalmente no se encuentran uniformemente distribuidas, los tratamientos que estos requieren se aplican de manera homogénea, por lo que proponen el uso de los RPAS como una solución potencialmente idónea a corto plazo ya que su versatilidad, en términos de avanzar, cambiar de dirección o permanecer en el aire en una posición fija, permiten adquirir imágenes de la tierra en casi cualquier condición y a partir de estas llevar a cabo los análisis pertinentes.

Sin embargo, Montesinos Aranda (2015) a pesar de reconocer las aplicaciones potenciales de los drones en la agricultura de precisión, también indica carencias que tienen los RPAS respecto a otras plataformas aéreas como las satélites o los aviones tripulados. Principalmente, hace referencia a la menor utilidad de los drones en la adquisición de datos en grandes superficies de plantaciones, por ejemplo, olivares o viñedos, debido a que para estas superficies es más útil adquirir las imágenes aéreas procedentes de los satélites, ya que en una sola imagen se recubre una superficie de 10.000 ha. Otra de las desventajas de los drones respecto a imágenes satelitales es la menor capacidad respecto a su resolución espectral (Díaz García-Cervigón, 2015), teniendo menores resoluciones espectrales las cámaras con las que están trabajando los drones respecto a las utilizadas en los satélites.

De esta manera se determina que, tanto los drones como las imágenes satelitales, poseen ventajas y desventajas al momento de ser aplicados para llevar a cabo una agricultura de precisión; por un lado, los drones, siempre y cuando se les realice un procesamiento adecuado, poseen una mayor resolución espacial que brinda pixeles de pequeño tamaño de la vegetación, tienen mayor resolución

temporal al poder capturar imágenes en días nublados (Barreiro & Valero, 2014), y, además, no requieren de calibraciones atmosféricas ya que las imágenes son captadas a una altura por debajo de las nubes y no se ven afectadas por la interferencia de la atmósfera (Hemav, 2018), sin embargo, poseen una mala cobertura ya que el área que pueden captar es muy limitada (Berni et al., 2009), si las condiciones meteorológicas no son las ideales no se puede volar y su escasa capacidad de carga los restringe al uso de cámaras RGB o multiespectrales y en pocos casos hiperespectrales como las empleadas por plataformas más robustas (Barreiro & Valero, 2014), en general, el costo de usarlos es mucho más elevado en comparación con las imágenes de satélite que se pueden encontrar de forma gratuita, aunado al hecho de que en algunos países se requieren permisos especiales para volar y los pilotos deben trasladarse junto con el equipo a la zona de estudio y estar completamente capacitados para que este vuele de la mejor manera (Auravant, 2018).

Por otro lado, las imágenes satelitales disponen de una elevada resolución espectral (Barreiro & Valero, 2014), es decir, el número de bandas en que captan la radiación reflejada es mayor, y el área que abarcan las imágenes es muy extensa (Paz Cerda, 2004), no obstante, los datos que se obtienen desde el satélite se ven afectados por la distancia a la que son capturados (Morales Cruz, 2013), las diferencias de reflectancia del suelo, ya sea debido a las presencia de nubes, sombras, impurezas o a su tipología, o por la propia vegetación objeto de estudio, puede inducir a un tratamiento de datos erróneos en comparación con un sistema de más precisión, razón por la cual requieren ser calibradas mediante correcciones geométricas y atmosféricas, y estas últimas, pueden ocasionar distorsiones entre imágenes adquiridas en tiempos diferentes debido a una concentración de aerosoles presentes en la atmósfera (Hemav, 2018).

Montesinos et al., (2012) pone en evidencia la necesidad de complementar la utilización de tecnologías de RPAS integrándolas junto con otros tipos de sensores, debido a las limitaciones que pueden tener los drones sobre todo en lo relacionado con la capacidad de cubrir grandes superficies, que al fin y al cabo son los clientes potenciales en cuanto a demanda de agricultura de precisión.

Algunos estudios han tratado de desarrollar métodos que vinculen información de RPAS e imágenes satelitales para proporcionar un producto de datos mejorado con una utilidad más allá de lo generado por cada componente. Se plantea que los datos resultantes del RPAS se utilicen como herramienta que proporcione información equivalente a la verdad-terreno y como posterior validación de información de satélite (Pla et al., 2017), con el fin de que la combinación de los datos espectrales de una imagen satelital con la información captada por el RPAS que posibilita imágenes de mayor resolución espacial, permita el uso de más información que puede mejorar las clasificaciones logrando que la imagen general del terreno sea mucho más precisa y permitiendo que se pueda automatizar el proceso.

Gevaert et al., (2015) y Dash et al., (2018) prueban la sensibilidad de las imágenes multiespectrales de series temporales obtenidas de RPAS y la información hiperespectral de imágenes satelitales. Dash et al. realizan este procedimiento con el fin detectar estrés por herbicidas en una especie arbórea, obteniendo que los datos de RPAS detectan el estrés hasta el nivel de árboles individuales, mientras que los datos del satélite Formosat 2 solo lo detectan en grupos de cuatro o más árboles; remuestran las imágenes de RPAS a la misma resolución espacial de las imágenes satelitales y revelan que las diferencias de sensibilidad no se dan únicamente como consecuencia de una resolución espacial diferente, por lo que hacen uso de índices de vegetación para optimizar la

detección del estrés. La relación entre los datos de RPAS y los índices espectrales derivados del satélite es explorada mediante regresión de modelos como mínimos cuadrados ordinarios y Random Forest. Con el fin de proponer una solución a las limitantes de ambas imágenes de manera individual, Gevaert et al. propone la fusión de imágenes de ambos sensores a través métodos que puedan retener la información hiperespectral de las imágenes satelitales RapidEye al usar la información adicional que proveen las observaciones multispectrales de RPAS y así, obtener un perfil temporal más completo; la metodología propuesta para obtener un STRS (Spectral-temporal response surfaces) se basa en la teoría bayesiana que permite cuantificar las incertidumbres, a partir de la cual obtienen un correlación de 0.953 con respecto a las mediciones espectrales realizadas en campo.

De igual forma, Gray et al., (2018) usan imágenes de RPAS como herramienta para entrenar al algoritmo de aprendizaje supervisado Support Vector Machine y clasificar las imágenes satelitales de WorldView-3 y RapidEye de la Reserva Rachel Carson en el Norte Carolina, EE. UU. Emplearon evaluaciones de precisión de RPAS y de campo para la comparación entre métodos de validación. Crearon y examinaron una serie de índices y capas, incluyendo la textura, NDVI, y un DEM LIDAR. El resultado demuestra la precisión de la clasificación a la par con los resultados de trabajo de campo (93% RPAS y 93% de campo para WorldView-3; 92% RPAS y 87% de campo para RapidEye). Tanto WorldView-3 como RapidEye se consideraron valiosas fuentes de imágenes para la clasificación de hábitats, concluyen que la información de RPAS puede ser altamente efectiva entrenando y validando las imágenes de satélite.

La combinación de imágenes satelitales Landsat y de RPAS es hecha por autores como Jenerowicz & Woroszkiewicz, (2016), quienes llevan a cabo la fusión de las imágenes para aplicaciones agrícolas, especialmente para distinguir tipos de cultivos, donde describen enfoques de pansharpening como métodos de fusión de datos elegidos para imágenes satelitales y datos obtenidos de altitudes bajas; y por Szantoi et al., (2017), donde dicha combinación es realizada con el fin de monitorear el hábitat de los orangutanes, los datos del RPAS se usan para entrenar al algoritmo de clasificación Random Forest y se aplican a los datos de Landsat para diferenciar clases de uso de la tierra (plantación de palma aceitera, reforestación, bosque talado) relevantes para la conservación del hábitat del orangután, alcanzan una precisión considerada como exitosa al ser esta superior al 75%; en ambos casos se concluye que la combinación de las imágenes da como resultado un producto con una mejor resolución espacial, temporal, y espectral, proporcionando una herramienta eficaz para discriminar clases de cobertura.

Por otro lado, Zhao et al., (2019) también llevan a cabo un procedimiento de fusión de imágenes satelitales con imágenes de RPAS, pero en este caso con imágenes Sentinel-2A para mapear la distribución de cultivos a una escala espacial más fina. Hacen uso de estas imágenes con un nivel de procesamiento 1C ya que incluye correcciones radiométricas y geométricas con precisión de subpíxel. Mediante Random Forest, Support Vector Machine y Neural Net, clasifican tanto las imágenes originales como las fusionadas en diez categorías: arroz, maíz, soja, trigo sarraceno, otras vegetaciones, invernaderos, tierra descubierta, agua, carreteras y casas. La fusión es realizada con el método de Gram-Schmidt debido a que es rápido y fácil de implementar y genera imágenes fusionadas con alta calidad de integración de color y detalles espaciales, la realizan con una serie de imágenes de RPAS a diferentes resoluciones, 0.03m, 0.10m, 0.50m, 1.00m, 3.00m, los mejores

resultados son obtenidos con el algoritmo de clasificación Random Forest y a partir de la fusión de la imagen de 0.10m.

Como conclusión, los autores concuerdan en que la posibilidad de combinar datos satelitales con información de RPAS puede permitir realizar la clasificación de coberturas con mayor precisión a lo largo de las dimensiones espaciales, espectrales y temporales, en aplicaciones como la agricultura de precisión, y así, se podría cerrar la brecha entre los requerimientos de los datos y las capacidades de los sensores.

Para llevar a cabo el procesamiento de imágenes que contienen información espectral, existen diversos softwares que permiten llevar a cabo dicha tarea, entre ellos, GEE - Google Earth Engine; esta plataforma que, además, permite hacer uso del catálogo de imágenes satelitales de Google, implementa una interfaz de usuario amigable que utiliza herramientas de procesamiento y análisis espacial con ayuda de algoritmos y líneas de códigos en lenguaje JavaScript. Dicha herramienta permite desarrollar procesos complejos que involucran gran cantidad de información dentro de los servidores de Google y finalmente permite visualizar los resultados e inclusive descárgalos para ser utilizados posteriormente ya sea dentro de la misma herramienta o en herramientas externas de análisis espacial como cualquier Sistema de Información Geográfico (Google, 2020).

Para llevar a cabo la clasificación, el algoritmo Random Forest ha sido utilizado en diferentes proyectos debido a la precisión de sus resultados, lo que demuestra la viabilidad de utilizar este método para los fines del proyecto. Belgiu & Drăgu, (2016) realizan una revisión de las aplicaciones de este algoritmo en la teledetección, donde encuentran que, como consecuencia de la gran cantidad de árboles de decisión producidos al seleccionar aleatoriamente un subconjunto de muestras de entrenamiento y un subconjunto de variables para dividir cada nodo del árbol, este clasificador es menos sensible a la calidad de las muestras de entrenamiento y al sobreajuste en comparación con otros clasificadores de machine learning. Además, a partir de las investigaciones revisadas, establecen que la recomendación más común es que el número de árboles empleados debe ser 500 y el número de ramificaciones debe ser la raíz cuadrada de las coberturas de interés.

Por ejemplo, Zurqani et al., (2017) realizan una clasificación de una colección multitemporal de imágenes Landsat mediante GEE y emplean Random Forest como clasificador, obtienen resultados catalogados como exitosos, concluyendo además que el clasificador tiene como limitante la necesidad de datos de referencia provenientes de imágenes de alta resolución espacial para trabajar con un set de entrenamiento de calidad, por lo que para realizar clasificaciones que datan de fechas antiguas se deben evaluar otras alternativas. De igual manera, ambas herramientas, la plataforma GEE y el clasificador Random Forest, son utilizados por Teluguntla et al., (2018) y por Agarwal & Nagendra, (2020), para llevar a cabo la clasificación supervisada por pixel de las zonas de cultivos y pastos en toda Australia y China y para monitorear la expansión urbana de las diez ciudades más pobladas de la India, respectivamente. Teluguntla et al., a partir de información de campo e interpretación de imágenes de muy alta resolución espacial (VHR), obtienen una clasificación general de más del 94%. Los resultados de Agarwal & Nagendr arrojaron índices de calidad cercanos al 80%.

Así pues, como consecuencia de que la revisión del estado del arte permite concluir que ningún conjunto de datos posibilita por sí solo la obtención de resultados óptimos, se plantea como solución, proponer una metodología de clasificación de imágenes de satélite que sea auxiliada por

los datos obtenidos de una imagen de mejor resolución espacial como las obtenidas mediante RPAS y así mejorar los procedimientos automatizados de clasificación al introducir un conjunto de parámetros para la valoración de la clasificación más allá de los valores digitales originales de los píxeles. Ambas tecnologías permiten obtener información valiosa, y, definidos los parámetros que condicionan el uso de una metodología u otra en función de los análisis que se requieran realizar sobre los cultivos, la combinación de estas supone el tener información útil o no de la vegetación.

4.2 Desarrollo de la metodología y el algoritmo de clasificación

A partir de los conocimientos generados de la revisión de la literatura y los conocimientos propios, se contribuyó al desarrollo de la metodología para llevar a cabo la integración de la imagen satelital Sentinel-2 y la imagen capturada mediante el RPAS a través de las siguientes actividades:

- Clasificación de conglomerados del departamento del Huila

Con el fin de definir algunos parámetros del script V03-01, correspondiente a los procesos de clasificación únicamente para una imagen satelital, se realizaron algunas pruebas para los conglomerados existentes en diferentes municipios del departamento del Huila; esto como consecuencia de que dicho departamento poseía el mayor error de definición de porcentajes de coberturas clasificadas mediante fotointerpretación, comparados con los resultados obtenidos en campo a través de la ENA.

De 37 municipios definidos por el equipo de trabajo, con un total de 52.738 Ha, se llevó a cabo la clasificación de 12 de ellos en la plataforma de Google Earth Engine. La Tabla 3 relaciona los municipios, el número de conglomerados existentes y el área clasificada en cada municipio.

Tabla 3. Conglomerados clasificados del departamento del Huila

Municipio	Número de conglomerados	Área clasificada
AIPE	5	3552,43
ALTAMIRA	3	1279,42
BARAYA	1	57,88
CAMPOALEGRE	14	2599,41
LA PLATA	23	2866,10
NEIVA	7	2107,04
PALERMO	15	6477,57
RIVERA	14	2075,53
SANTA MARÍA	5	616,54
TERUEL	4	369,63

TESALIA	4	1462,17
VILLAVIEJA	1	186,869
TOTAL	96	23.650,589

- Análisis estadístico de los resultados de la clasificación

Se llevó a cabo un análisis estadístico de los resultados obtenidos al clasificar todos los conglomerados existentes en el departamento del Huila. Teniendo en cuenta que el proceso se llevó a cabo mediante los algoritmos de clasificación *Random Forest* y *Support Vector Machine*, se analizaron los siguientes resultados:

- Tamaño y precisión de la muestra
- Precisión temática
- Rendimiento de la plataforma GEE
- Resultados de la clasificación
- Porcentaje de coincidencia con los resultados de la ENA

A partir de los análisis realizados para cada uno de los resultados anteriores, fue posible definir los siguientes parámetros:

- El algoritmo a utilizar en el proceso de clasificación en todas las versiones del Script: *Random Forest*.
- El número de árboles más adecuado para realizar la clasificación mediante *Random Forest*.
- Las clases de cobertura que son posibles y adecuadas identificar mediante el script y en la plataforma GEE.
- El número de muestras que se deben tomar por cobertura: aproximadamente 20 muestras.
- El tipo de geometría a utilizar en la selección de las muestras por cobertura.

4.2.1 Resultados

A través de las pruebas del script V03-01 y el planteamiento de mejoras del mismo, se consolidaron los pasos a realizar para integrar de forma efectiva los insumos y se identificaron los actores del proceso, obteniendo el flujo metodológico para el script V03-03, correspondiente a los procesos que permiten realizar la clasificación a partir de la imagen satelital integrada con las imagen de RPAS.

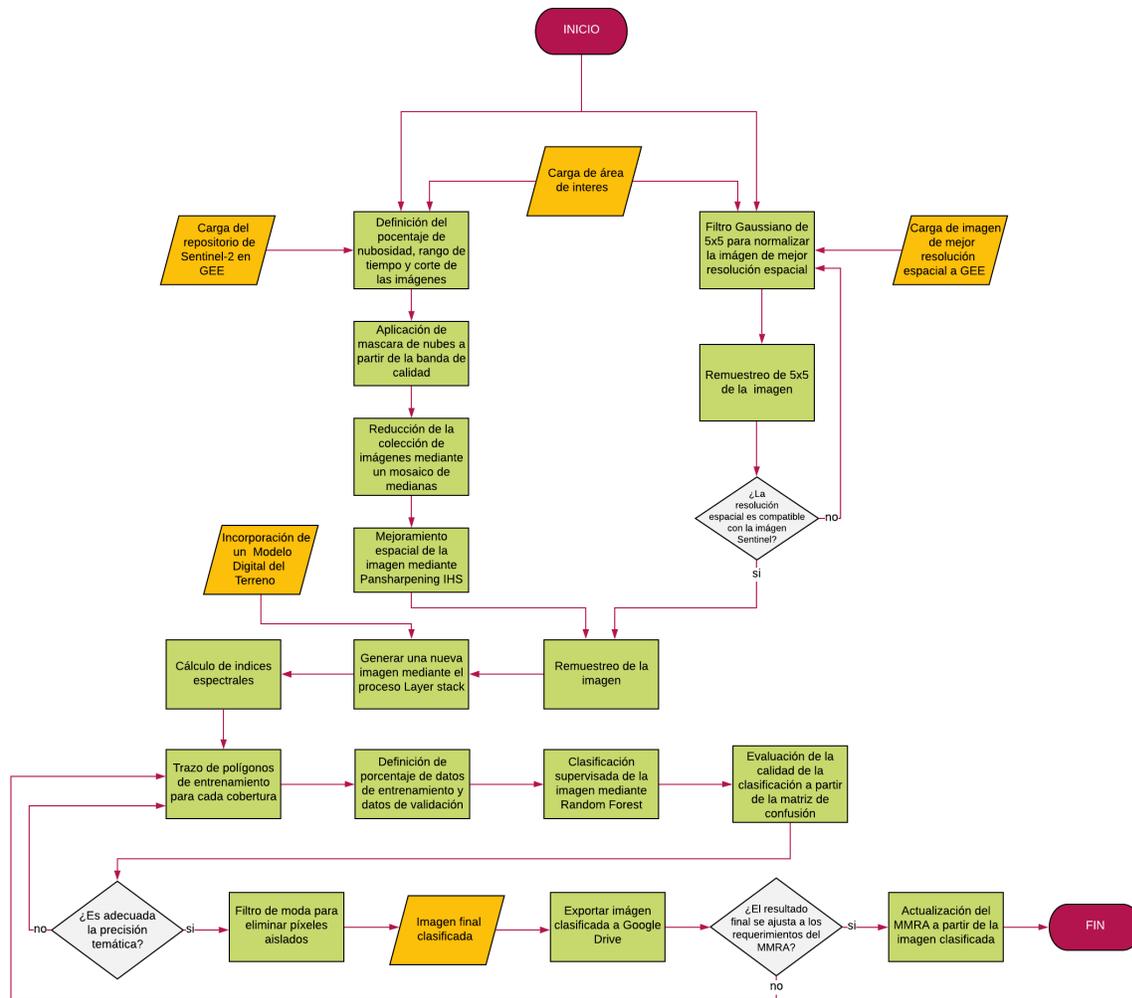


Ilustración 4. Esquema conceptual del flujo de trabajo

Dentro del flujo definido en la Ilustración 4, se establece como primer paso el cargue del repositorio Sentinel-2 en procesamiento SR (*COPERNICUS/S2_SR*), realizando la acotación en los siguientes aspectos:

Porcentaje de nubosidad: bajo este aspecto se selecciona las imágenes con un porcentaje igual o menor al indicado por el intérprete. Se debe tener presente que esto proviene de la metadata de la imagen, es decir, es la concentración estimada que se encuentra presente en el total de la imagen.

Rango de tiempo: Indica la temporalidad en la que se desea realizar la búsqueda de imágenes.

Definición y corte a partir del área de estudio: este filtro define y realiza el corte de la zona de estudio en el que se desea realizar la búsqueda, correspondiendo a los conglomerados de interés del MMRA.

De forma adicional, es aplicada una máscara de nubosidad a cada imagen de la colección, que es producto de la identificación y corte de las zonas que, de acuerdo a la banda de calidad QA60, posee

presencia de nubes y cirrus (Traganos et al., 2018). Esta acción disminuye la posibilidad de que existan nubes en el mosaico que es utilizado como insumo de la clasificación.

Con el propósito de realizar la clasificación a partir de una sola imagen, es empleada una reducción de la colección de las imágenes, generando un mosaico de medianas, que consiste en seleccionar el valor de la mediana de cada pixel, esta técnica es plenamente utilizada porque, de acuerdo con Flood (2013) su uso genera una imagen homogénea en el espacio y tiempo que probablemente no posee afectaciones por presencia de nubes y fuentes de contaminación.

Debido a que se encuentra diferencias en la resolución espacial en las bandas espectrales RGB con las de *SWIR-1*, *SWIR-2* y *Red Edge* de Sentinel-2, puesto que esta varía de 10 m a 20 m, es necesario realizarle al mosaico de medianas el proceso *pansharpening*, que consiste en un proceso de fusión de imágenes que busca mejorar la resolución espacial de las bandas sin perder la resolución espectral que poseen (Li et al., 2018). El proceso empleado se realiza a partir de la transformación IHS (Intensidad, Matiz y Saturación), el cual es frecuentemente utilizado en imágenes satelitales HR (Kaplan, 2018) y su implementación en GEE es sencilla (HC, 2019).

Con el objetivo de integrar las imágenes de mejor resolución espacial al mosaico de medianas, se consideraron las limitantes relacionadas con las resoluciones espaciales y temporales de cada uno de los insumos, debido a que no son comparables entre sí (Rocchini, 2007), esto como consecuencia de que el mosaico de medianas corresponde a un resumen del comportamiento espectral de las coberturas en un lapso de tiempo y tomado fuera de la atmósfera, mientras que la imagen de mejor resolución espacial representa la toma espectral de una fecha específica por debajo de la atmósfera a un tamaño de píxel mucho más detallado. Lo anterior obliga a que, para que existan puntos comunes de comparación, las dos imágenes se deben encontrar en un nivel de reflectancia preferiblemente cercano al de superficie (Nonni et al., 2018) y que para realizar la comparación espacial se siga un procedimiento estricto y exhausto de normalización.

Para llevar a cabo dicha normalización de la imagen de mejor resolución espacial, se emplea un filtro de textura gaussiano (Ariputhiran & S. Gandhimathi, 2013) de tamaño 5x5 y se remuestrea cinco veces su tamaño de píxel original de forma escalonada hasta que sea comparable espacialmente con el mosaico Sentinel-2 (Nonni et al., 2018) como se visualiza en la Ilustración 5, para de esta forma realizar un remuestreo ponderado, cabiendo mencionar, además, que la aplicación de este filtro recrea los fundamentos teóricos de toma del *Instantaneous Field Of View* (IFOV) (Anon n.d.). Para fijar el número de iteraciones necesarias se aplica la Fórmula 1, que está condicionada en que la resolución espacial objetivo final es de 10 m.

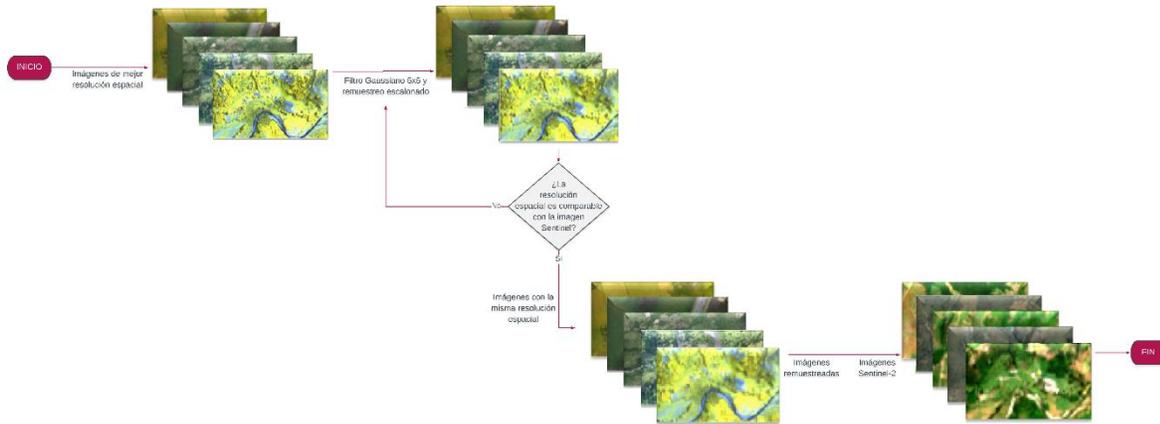


Ilustración 5. Remuestreo de la imagen de mejor resolución espacial

$$n = E \left[\frac{\ln \left(\frac{10}{\text{resolucion original}} \right)}{\ln 5} \right] \quad (1)$$

De acuerdo con Balzter et. al. (2015) una buena práctica para mejorar la diferenciación de las coberturas de interés empleando algoritmos de clasificación supervisados de *Machine Learning*, es utilizando información disponible de la topografía de la zona de estudio, en este caso se refiere a emplear un Modelo Digital de Terreno (MDT). Por la razón anterior, se agrega al mosaico de medianas el corte del MDT SRTM de resolución espacial de 30 m disponible en el repositorio de GEE y de forma adicional, a partir de este mismo, se calcula y agrega un ráster de pendientes (Chen et al., 2017).

Como información de referencia espectral para mejorar la clasificación final, tanto para el mosaico de medianas como para la imagen de mejor resolución espacial remuestreada, se calculan y agregan los índices espectrales de la Tabla 4. Se especifica que, de acuerdo con la resolución espectral de la imagen de mejor resolución espacial, se realiza el cálculo de los índices espectrales.

Tabla 4. Índices espectrales agregados

Nombre	Significado	Fórmula	Características	Fuente
NDVI	Índice de vegetación de diferencia normalizada	$\frac{NIR - RED}{NIR + RED}$	Se usa para llevar seguimiento de las condiciones de la cubierta vegetal y ayuda a diferenciar la vegetación de otras coberturas. Varía entre -1 a 1, los resultados cercanos a 1 indican la presencia de vegetación, los resultados negativos o cercanos a cero indican la presencia de cuerpos de agua, nubes, nieve, zonas de suelo desnudo y rocas.	(Viana et al., 2019)

SIPI	Índice de pigmentación intensiva de la estructura	$\frac{NIR - BLUE}{NIR - RED}$	Utilizado para monitorear la salud de las plantas en regiones con alta variabilidad en la estructura del dosel o el índice de área de la hoja. Estima la proporción de carotenoides a clorofila. Varía entre 0 a 2, donde la vegetación verde saludable es de 0.8 a 1.8	EOS, 2020
PVR	Relación de vigor fotosintético	$\frac{GREEN - RED}{GREEN + RED}$	Se usa para detectar variaciones relativas en la densidad de cultivos. Los valores altos en los resultados indican vegetación sana y fotosintéticamente activa, los valores más bajos se registran para el dosel amarillo, fotosintéticamente débil debido a factores como la deficiencia de nutrientes, el estrés o el inicio de la senescencia.	(Metternicht, 2003)
VARI	Índice de resistencia atmosféricamente visible	$\frac{GREEN - RED}{GREEN + RED - BLUE}$	Diseñado para datos RGB. Este índice es mínimamente sensible a los efectos atmosféricos favoreciendo la estimación de la fracción de vegetación. Se usa para analizar los cultivos en todas las etapas de crecimiento.	(Gitelson et al., 2002)
GARI	Índice Verde Resistente Atmosféricamente	$\frac{NIR - (GREEN - \lambda(BLUE - RED))}{NIR + (GREEN - \lambda(BLUE - RED))}$ $\lambda \rightarrow \text{parámetro que controla la corrección atmosférica}$	Se usa para determinar el contenido en clorofila aplicando un parámetro λ , el cual determina el tipo de aerosoles y minimiza la dispersión atmosférica.	(Caseras Saez, 2017)
NDBI	Índice de áreas construidas de diferencia normalizada	$\frac{SWIR - NIR}{SWIR + NIR}$	Se usa para la identificación de zonas edificadas y suelo desnudo, estos se ven indicados en los valores positivos, los valores cercanos a cero indican coberturas de bosque y agricultura y valores negativos indican agua.	(Viana et al., 2019)
NDWI	Índice de agua de diferencia normalizada	$\frac{GREEN - NIR}{GREEN + NIR}$	Se usa para medir la cantidad de agua que posee la vegetación o el nivel de saturación de humedad que posee el suelo. Varía entre -1 a 1, Valores positivos indican la presencia de agua, valores iguales a cero o negativos indican la ausencia de ella.	(Muñoz Aguayo, 2013)

El algoritmo que se utiliza para la clasificación supervisada basada en píxeles es de *Machine Learning, Random Forest*, ya que, según la revisión de estado del arte y los resultados obtenidos a partir de las clasificaciones realizadas en los conglomerados del departamento del Huila, es el más usado y con mejores resultados, obteniendo la mejor exactitud temática (Belgiu & Csillik, 2017; Chen et al., 2017; Zurqani et al., 2017).

La clasificación se realiza a partir de los polígonos de entrenamiento trazados manualmente por el intérprete para cada cobertura siguiendo el procedimiento mostrado en la Ilustración 6. La muestra es dividida en polígonos de entrenamiento y de validación. Para la selección del set de datos de validación, se toman las recomendaciones realizadas por Millard & Richardson (2015), en cuestión de que el 30% del total de los polígonos corresponde a la información de validación; su selección es a partir de números aleatorios con una semilla predispuesta, que en este caso es cero, y producto de la verificación de un “*spatial join*” entre los datos de entrenamiento y los preseleccionados de validación, con el propósito de que no exista autocorrelación espacial con la muestra de entrenamiento.



Ilustración 6. Procedimiento para el trazo de polígonos de entrenamiento

Para el proyecto, se seleccionan las coberturas de la Tabla 5 con sus correspondientes equivalencias entre el MMRA y la ENA.

Tabla 5. Coberturas definidas en GEE

GEE		MMRA		ENA	
COBERTURA	DESCRIPCIÓN	CLASIFICACIÓN DE COBERTURA	DESCRIPCIÓN	CLASIFICACIÓN DE COBERTURA	DESCRIPCIÓN
1. CultivoSin	Aquellos que cultivos que no se pueden indentificar en la imagen como trasintorios o permanentes.	PORCCULT_SIN_CLAS	Porcentaje área sembrada de cultivos sin clasificar	AREA_BAR_H	Área en barbecho
11. SueloEnPreparacion	Suelo en proceso de trabajo con el fin de crear las condiciones adecuadas para la siembra de un cultivo.			AREA_DESC_H	Área en descanso
2. Transitorio	Agricultura intensiva con especies como arroz, algodón, sorgo, maíz, frijol, papa, en todos los pisos térmicos, barbecho, cultivos confinados como horticultura, flores y viveros.	PORCTRANSITORIO	Porcentaje área sembrada de cultivos transitorios	AREA_TRAN_H	Área en transitorios

3. Permanente	Agricultura de caña para panela; Caña de azúcar; palma de aceite; plantaciones de banano y plátano; agricultura del café; plantaciones frutales en los diferentes pisos térmicos (Cacao, cítricos, vides, peras, manzanas)	PORCPERMANENTE	Porcentaje área sembrada de cultivos permanentes	AREA_PER_H	Área en permanentes
4. Forestales	Plantación de especies forestales, con una técnica de siembra específica	PORCFORRESTALES	Porcentaje área sembrada en plantaciones forestales	AREA_BOSPLA_H	Área en bosques plantados
5. Pastos	Pastos naturales y/o introducidos (Herbáceos o arbustivos); pastos con árboles y/o arbustos; pastos introducidos manejados; asociación de pastos y matorrales.	PORCPASTOS	Porcentaje área sembrada en pastos	AREA_PASTO_H	Área en pastos y forrajes
13. VegetacionSecundaria	Cobertura vegetal originada por el proceso de sucesión de la vegetación natural que se origina luego de la intervención o por la destrucción de la vegetación primaria, que puede encontrarse en recuperación tendiendo al estado original.	PORCOTROSUSOS	Porcentaje área en otros usos y coberturas	AREA_MALRAS_H	Área en malezas y rastrojos
6. OtrosUsos	Área en vegetación de sabana, área en tundras o páramos con vegetación herbácea o arbustiva, área en vegetación xerofítica (desierto), construcciones asociadas a la actividad pecuaria , ganadera y piscícola como corrales y galpones, eriales o suelos desnudos, erosión natural o provocadas, afloramiento rocoso.			AREA_VEGSAB_H	Área en vegetación de sabana
				AREA_VEGPAR_H	Área en vegetación de paramos
				AREA_VEGXER_H	Área en vegetación xerofítica (desierto)
				AREA_INFPEC_H	Área en infraestructura pecuaria
				AREA_INFPIH_H	Área en infraestructura piscícola
				AREA_INFAGR_H	Área en infraestructura ganadera
10. CuerposAgua	Cuerpos de agua natural, pantanos permanentes y temporales.			AREA_CUAGUA_H	Área en cuerpos de agua
12. BancoArena	Riveras y playas, dunas	-	-		
7. BosqueNatural	Bosque denso alto o bajo de tierra firme o inundable, bosque fragmentado con pastos y cultivos, bosques de galería.	PORCBOSQUENATURAL	Porcentaje área en bosque natural	AREA_BOSNAT_H	Área en bosques naturales
8. NoAgropecuario	Construcciones nucleadas o dispersas de caseríos, parques industriales, viviendas parques turísticos, bodegas, áreas mineras a cielo abierto, vías.	PORCNOAGROPECUARIO	Porcentaje área en predios no agropecuarios	AREA_OTFIN_H	Área en otros fines
9. Nubes	Fración de cielo cubierto por nubes en un lugar en particular.	-	-	-	-

Como indicadores de la calidad temática de la clasificación, se utilizan las métricas más representativas que corresponden a porcentaje de clasificación del mapa, calidad del productor, calidad del usuario y kappa, todas las anteriores son derivadas de la matriz de confusión generada a partir de la información de entrenamiento y por otro lado de validación. Foody (2002) define a la matriz de confusión como el resultado del cruce tabulado de las clases de la imagen y de la información de terreno o referencia. Por consiguiente, el porcentaje de clasificación del mapa

corresponde a la sumatoria de los píxeles con clasificación correcta (diagonal de la matriz de confusión) sobre el número total de píxeles (Liu et al., 2007) mientras que la calidad del productor y del usuario son derivadas de considerar los píxeles correctamente clasificados sobre el total de los píxeles considerados conformes de la información de referencia y sobre el resultado de la clasificación respectivamente (Foody, 2002). Por último, se reporta el índice kappa debido a que no posee sesgo derivado de los errores de omisión y comisión (Foody, 2002; Liu et al., 2007).

La clasificación final es seleccionada producto de la revisión exhausta de los indicadores de exactitud temática junto a una validación visual aleatoria de las coberturas clasificadas. A la clasificación final le es aplicado un filtro de moda de tamaño 3x3, debido a que este procedimiento atenúa los píxeles aislados que fueron clasificados de forma puntual y permite que se cumpla con criterios técnicos cartográficos como el área mínima cartografiable, proporcionando un producto final mucho más homogéneo debido a que elimina errores residuales del proceso de clasificación (Stuckens et al., 2000).

El producto final se exporta a Google Drive, correspondiente a toda la imagen clasificada con el fin de tratarla con el flujo de trabajo establecido por la DIG para actualizar el MMRA.

La actualización del MMRA se realiza a partir de los conglomerados existentes, estableciendo para cada uno el porcentaje correspondiente a cada cobertura de la **¡Error! No se encuentra el origen de la referencia.** Esta información es capturada a través del empleo de software GIS, debido a que GEE no es tan apropiado para el análisis de datos vectoriales.

Como soporte de la implementación del flujo propuesto, en la Tabla 6 se comparten los enlaces con los scripts que contienen total o parcialmente los procesos descritos en la Ilustración 4.

Tabla 6. Scripts de los procesos

Script	Enlace	Descripción
V03-01	https://code.earthengine.google.com/5d3a324bb88ab057b277688bdedcf08b	Posee los procesos que permiten realizar la clasificación solo a partir de la imagen satelital
V03-02	https://code.earthengine.google.com/145bd2d3c67dae3b4273065a663c654a	Posee los procesos que permiten realizar la clasificación solo a partir de la imagen satelital, sin el proceso de pansharpning
V03-03	https://code.earthengine.google.com/e47fb21921955866ea5479a6a35f4cb8	Posee todos los procesos

4.3 Clasificación de las imágenes mediante el script V03-03

Como consecuencia de que el aislamiento preventivo para el evitar el contagio del Covid-19 no permitió salir a campo y tomar las imágenes de los cultivos de frijol y papa en sus diferentes estados fenológicos, como se tenía planteado al inicio del proyecto, fue necesario buscar imágenes provenientes de otras fuentes. En esta sección, se muestran los resultados obtenidos al clasificar las imágenes correspondientes a los siguientes municipios:

- Salento, Quindío
- Tununguá, Boyacá
- Pajarito, Boyacá
- Cómbita, Boyacá
- Cota, Cundinamarca.

Las pruebas se realizaron con el Script V03-01, que contiene los procesos que permiten realizar la clasificación solo a partir de la imagen satelital, y con el Script V03-03, que posee los procesos que permiten realizar la clasificación a partir de la integración de la imagen satelital con la imagen de RPAS, esto con el fin de comparar los resultados y definir las cualidades de cada uno para zonas con diferentes características.

Se consolidaron y analizaron los siguientes resultados de todas las imágenes utilizadas en el proyecto y a partir de las clasificaciones realizadas con ambas versiones del script:

- Indicadores de calidad temática
 - Calidad de los datos de entrenamiento
 - Calidad de los datos de validación
 - Índices de calidad temática (Productor, consumidor, error de omisión, error de comisión, precisión general, índice kappa)
 - Áreas clasificadas

4.3.1 Resultados

4.3.1.1 Pruebas realizadas en Salento

Las clasificaciones realizadas en esta zona se llevaron a cabo para cultivos agregados y para cultivos desagregados, ya que al realizar la identificación de las coberturas presentes en la imagen de RPAS se llegó a nivel de cultivo discriminado, por lo tanto, es posible evaluar los alcances de los algoritmos si se cuenta con este tipo de información.

- Producto del proceso de remuestreo

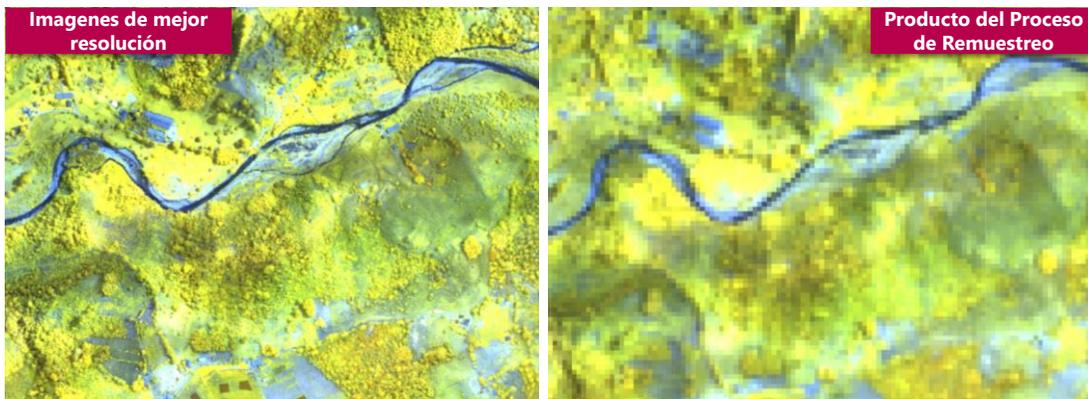


Ilustración 7. Proceso de remuestreo de la imagen de Salento - Combinación B4,B3,B2

- Pruebas para cultivos agregados
 - Muestras de entrenamiento

Las muestras de entrenamiento se capturaron tomando como base la interpretación realizada por un experto del grupo de MMRA. La Tabla 7 indica el número de muestras usadas por cada cobertura.

Tabla 7. Muestras de entrenamiento en cultivos agregados para Salento

Cobertura	Número de muestras	Cobertura	Número de muestras
1. Cultivo sin	0	8. No Agropecuario	44
2. Transitorio	43	9. Nubes	0
3. Permanente	267	10. Cuerpos de Agua	14
4. Forestales	0	11. Suelo en Preparación	0
5. Pastos	73	12. Banco de Arena	24
6. Otros Usos	0	13. Vegetación Secundaria	0
7. Bosque Natural	47	14. Sombras	0

Tabla 8. Índices de calidad temática en cultivos agregados para la zona de Salento

Categoría	Script V03-01				Script V03-03			
	Productor	Error de Omisión	Consumidor	Error de comisión	Productor	Error de omisión	Consumidor	Error de comisión
Cultivo sin	0	0	0	0	0	0	0	0
Transitorio	0,902	9,8%	0,529	47,1%	0,976	2,4%	0,769	23,1%
Permanente	0,911	8,9%	0,899	10,1%	0,939	6,1%	0,905	9,5%
Forestales	0	0	0	0	0	0	0	0
Pastos	0,721	27,9%	0,98	2,0%	0,706	29,4%	0,96	4,0%
Otros Usos	0	0	0	0	0	0	0	0
Bosque Natural	0,984	1,6%	0,988	1,2%	0,992	0,8%	0,99	1,0%
No Agropecuario	0,531	46,9%	1	0,0%	0,688	31,2%	0,957	4,3%
Nubes	0	0	0	0	0	0	0	0
Cuerpos de Agua	0,909	9,1%	0,938	6,2%	0,939	6,1%	0,861	13,9%
Suelo en Preparación	0	0	0	0	0	0	0	0
Banco de Arena	0,889	11,1%	0,667	33,3%	0,889	11,1%	0,8	20,0%
Vegetación Secundaria	0	0	0	0	0	0	0	0
Sombras	0	0	0	0	0	0	0	0
Precisión General	0,922				0,942			
Indicador Kappa	0,878				0,909			

- Resultados de la clasificación

Al comparar las áreas de las clasificaciones obtenidas a partir de los dos scripts, Tabla 9, se puede observar que en el script V03-03 aumento significativamente el área para los cultivos permanentes, por otro lado, principalmente se presenta una disminución en las áreas para pastos, bosque natural y banco de arena.

Tabla 9. Áreas por cobertura en cultivos agregados de Salento

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
1. Cultivo sin	0	1. Cultivo sin	0
2. Transitorio	22,35	2. Transitorio	17,29
3. Permanente	1662,82	3. Permanente	1719,67
4. Forestales	0	4. Forestales	0
5. Pastos	194,23	5. Pastos	167,81
6. Otros Usos	0	6. Otros Usos	0
7. Bosque Natural	1866,71	7. Bosque Natural	1849,81
8. No Agropecuario	24,42	8. No Agropecuario	26,17
9. Nubes	0	9. Nubes	0,00
10. Cuerpos de Agua	24,90	10. Cuerpos de Agua	20,38
11. Suelo en preparación	0	11. Suelo en preparación	0,00
12. Banco arena	10,66	12. Banco arena	4,95
13. Vegetación Secundaria	0	13. Vegetación Secundaria	0
14. Sombras	0	14. Sombras	0
Total	3806,085	Total	3806,085

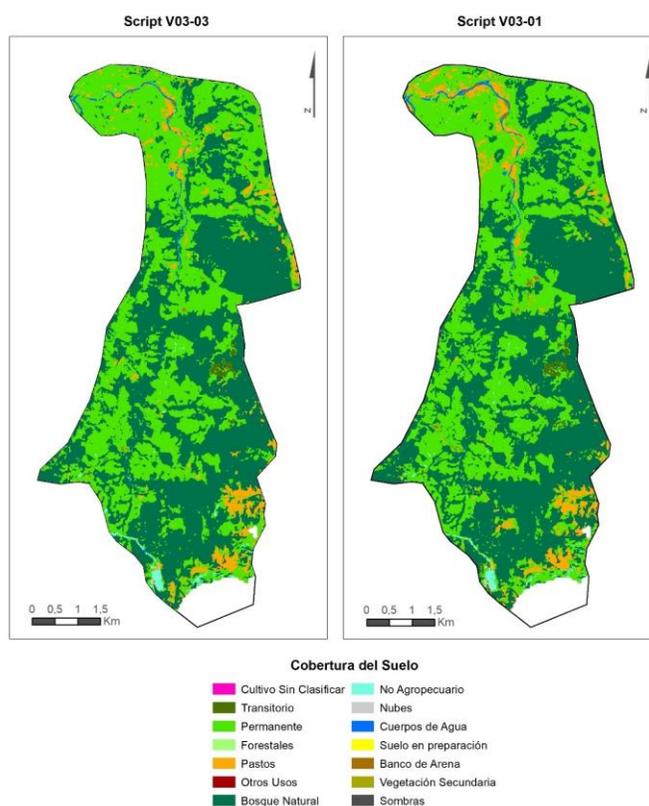


Ilustración 8. Resultado general de la clasificación en Salento con cultivos agregados

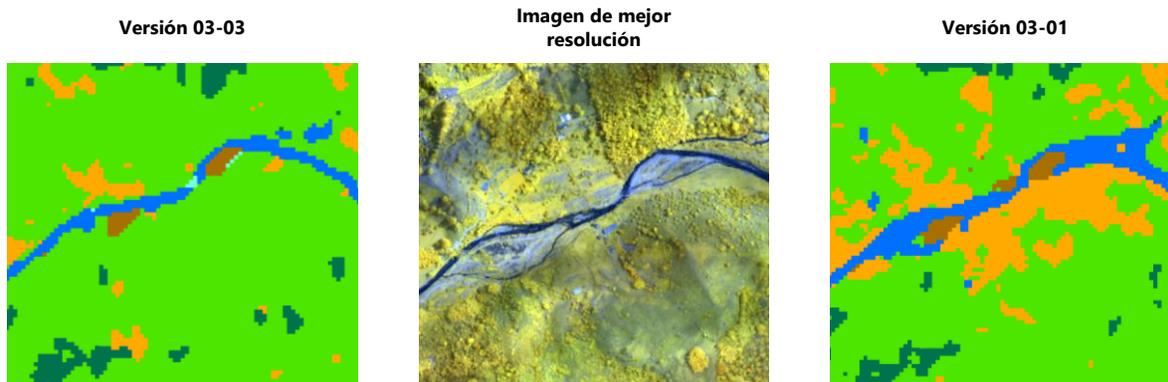


Ilustración 9. Resultados detallados en Salento

En la Ilustración 9, se evidencia como el cuerpo de agua, en color azul, clasificado con el script versión 03-03, es mucho más consistente con la realidad mostrada en la imagen de mejor resolución, en comparación con la clasificación realizada con la versión 03-01. En cambio, los pastos se muestran más consistente en la clasificación realizada con el script versión 03-01.

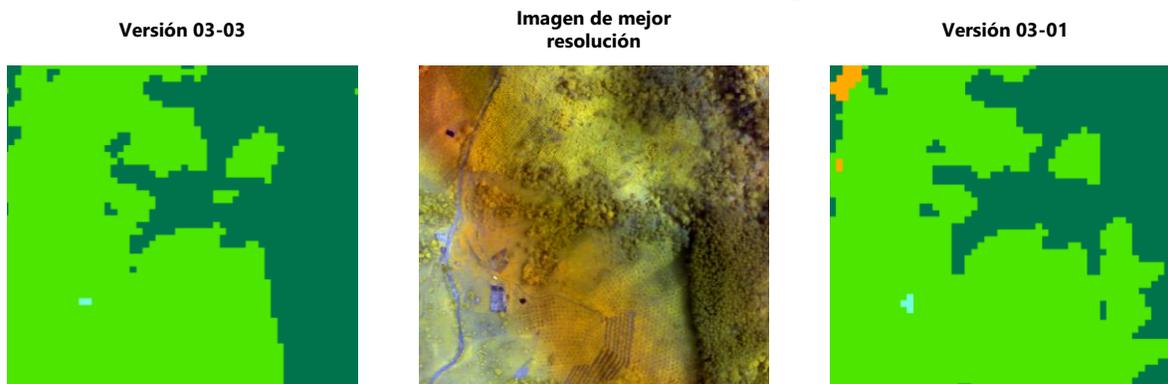


Ilustración 10. Resultados detallados en Salento

En la Ilustración 10, en la clasificación realizada con el script versión 03-03, se obtuvo una mejor definición del límite de entre los cultivos permanentes, mostrados en color verde claro, y los bosques naturales, mostrados en color verde oscuro.

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01
<https://code.earthengine.google.com/38a7c2490fa694d644be9ca5ca55fff0>
- V03-03
<https://code.earthengine.google.com/53504811effabd9665df20445a8a2a3d>

- Pruebas para cultivos desagregados
 - Muestras de entrenamiento

Tabla 10. Muestras de entrenamiento en cultivos desagregados para Salento

Cobertura	Número de muestras	Cobertura	Número de muestras
1. Aguacate	54	8. Construido	44
2. Transitorio	43	9. Nubes	0
3. Café con sombrero	36	10. Cuerpos de agua	14
4. Cítricos	22	11. Plátano	21
5. Pastos	73	12. Banco de arena	24
6. Café	91	13. Café con Barrera	43
7. Bosque natural	47	14. Sombras	0

Tabla 11. Índices de calidad de la clasificación en cultivos desagregados de Salento

Categoría	Script V03-01				Script V03-03			
	Productor	Error de Omisión	Consumidor	Error de comisión	Productor	Error de omisión	Consumidor	Error de comisión
Aguacate	0,75	25,0%	0,679	32,1%	0,829	17,1%	0,818	18,2%
Transitorio	0,951	4,9%	0,709	29,1%	0,951	4,9%	0,848	15,2%
Café con sombrero	0,925	7,5%	0,961	3,9%	0,943	5,7%	0,98	2,0%
Cítricos	0,714	28,6%	0,769	23,1%	0,714	28,6%	0,909	9,1%
Pastos	0,75	25,0%	0,586	41,4%	0,838	16,2%	0,613	38,7%
café	0,479	52,1%	0,672	32,8%	0,539	46,1%	0,703	29,7%
BosqueNatural	0,986	1,4%	0,962	3,8%	0,982	1,8%	0,96	4,0%
Construido	0,594	40,6%	0,905	9,5%	0,688	31,2%	0,917	8,3%
Nubes	0	0	0	0	0	0	0	0
CuerposAgua	0,939	6,1%	0,756	24,4%	0,939	6,1%	0,674	32,6%
platanos	0,75	25,0%	1	0,0%	0,917	8,3%	0,917	8,3%
BancoArena	0,889	11,1%	0,667	33,3%	0,889	11,1%	0,667	33,3%
Cafes con barreras	0,931	6,9%	0,92	8,0%	0,897	10,3%	0,963	3,7%
Sombras		0		0	0	0	0	0
Precisión General					0.848			
Indicador Kappa					0.797			
					0.869			
					0.825			

- Resultados de la clasificación

Al comparar las áreas de las clasificaciones obtenidas a partir de los dos scripts, Tabla 12, se puede observar que en el script V03-03 aumento significativamente el área para los cultivos de aguacate y café, por otra parte, se presenta una disminución en las áreas para pastos y bosque natural principalmente.

Tabla 12. Áreas por cobertura en cultivos agregados de Salento

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
1. Aguacate	464,76	1. Aguacate	602,43
2. Transitorio	32,83	2. Transitorio	22,12

3. Café con sombrío	187,42	3. Café con sombrío	170,8
4. Cítricos	37,79	4. Cítricos	24,12
5. Pastos	439,19	5. Pastos	332,74
6. Café	452,14	6. Café	527,2
7. BosqueNatural	2037,86	7. BosqueNatural	1990,63
8. Construido	34,65	8. Construido	33,32
9. Nubes	0	9. Nubes	0
10. CuerposAgua	28,88	10. Cuerpos Agua	28,6
11. Plátano	4,71	11. Plátano	4,63
12. BancoArena	12,64	12. BancoArena	10,86
13. Cafés con barreras	73,21	13. Cafés con barreras	58,64
14. Sombras	0	14. Sombras	0
Total	3806,085	Total	3806,085

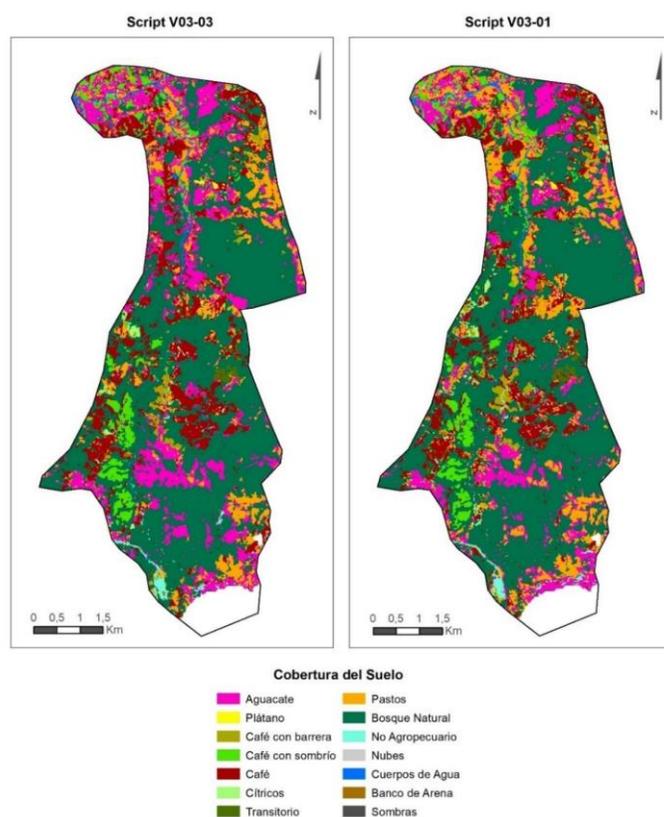


Ilustración 11. Resultado general de la clasificación en Salento con cultivos desagregados

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01
<https://code.earthengine.google.com/b716125362000d106544aeda8edea99e>
- V03-03
<https://code.earthengine.google.com/f8b6fef783a28cb6eb5914f8cdda4b9a>

4.3.1.2 Pruebas realizadas en Tununguá

- Producto del proceso de remuestreo



Ilustración 12. Proceso de remuestreo de la imagen de Tununguá - Combinación RGB

- Muestras de entrenamiento

Las muestras de entrenamiento se capturaron tomando como base la interpretación realizada por un experto del grupo del MMRA. La Tabla 13 indica el número de muestras usadas por cada cobertura.

Tabla 13. Muestras de entrenamiento para Tununguá

Cobertura	Número de muestras	Cobertura	Número de muestras
1. Cultivo sin clasificar	29	8. No Agropecuario	17
2. Transitorio	39	9. Nubes	11
3. Permanente	104	10. Cuerpos de Agua	0
4. Forestales	0	11. Suelo en preparación	0
5. Pastos	92	12. Banco arena	11
6. Otros Usos	31	13. Vegetación Secundaria	65
7. Bosque Natural	35	14. Sombras	0

Tabla 14. Índices de calidad de la clasificación de Tununguá

Categoría	Script V03-01				Script V03-03			
	Productor	Error de Omisión	Consumidor	Error de comisión	Productor	Error de omisión	Consumidor	Error de comisión
Cultivo sin	0,963	0,037	0,975	0,025	0,963	0,037	0,963	0,037
Transitorio	0,696	0,304	0,814	0,186	0,754	0,246	0,897	0,103
Permanente	0,736	0,264	0,688	0,312	0,747	0,253	0,701	0,299
Forestales	0	0	0	0	0	0	0	0
Pastos	0,804	0,196	0,789	0,211	0,822	0,178	0,784	0,216
Otros Usos	0,967	0,033	1,000	0,000	0,967	0,033	1,000	0
Bosque Natural	0,960	0,040	0,955	0,045	1,000	0	0,962	0,038
No Agropecuario	0,667	0,333	0,111	0,889	1,000	0	0,750	0,250
Nubes	0,766	0,234	0,990	0,010	0,855	0,145	1,000	0
Cuerpos de Agua	0	0	0	0	0	0	0	0
Suelo en preparación	0	0	0	0	0	0	0	0
Banco arena	1,000	0	1,000	0	1,000	0,000	1,000	0
Vegetación Secundaria	0,698	0,302	0,698	0,302	0,684	0,316	0,716	0,284
Sombras	0	0	0	0	0	0	0	0
Precisión General				0,800				0,823
Indicador Kappa				0,761				0,787

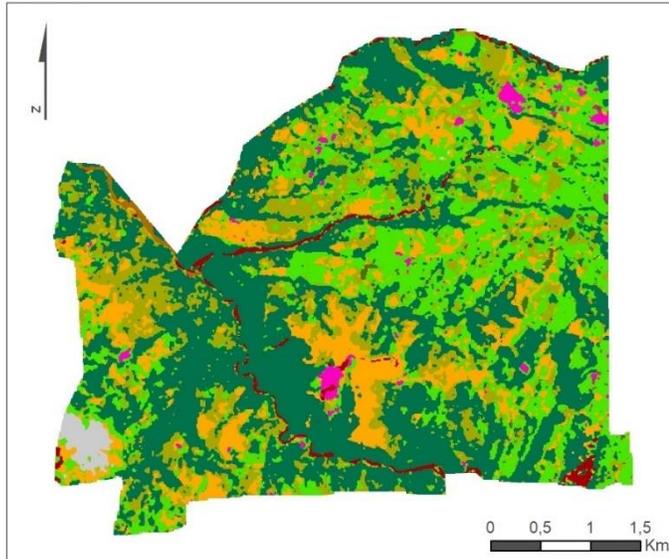
- Resultados de la clasificación

La Tabla 15, muestra la comparación de las áreas obtenidas de las clasificaciones a partir de los dos scripts, donde se puede observar que, en el script V03-03, aumentó el área para los cultivos permanentes y bosque natural, mientras que, se presenta una disminución en las áreas para vegetación secundaria y pastos.

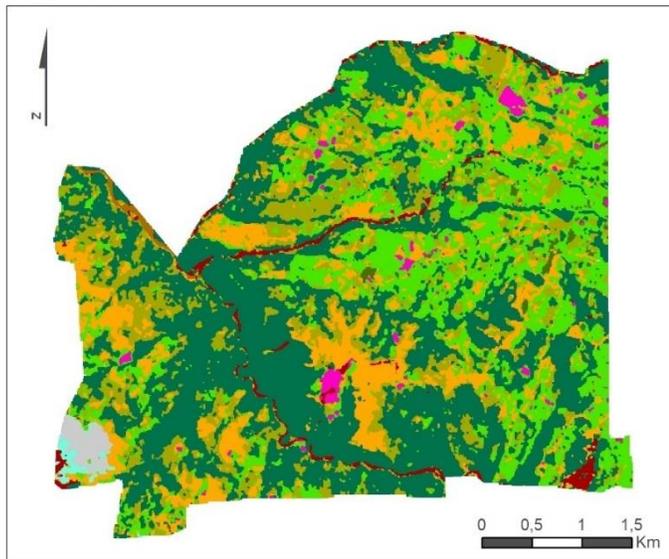
Tabla 15. Áreas por cobertura de Tununguá

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
1. Cultivo sin	26,989	1. Cultivo sin	24,477
2. Transitorio	17,524	2. Transitorio	12,726
3. Permanente	478,167	3. Permanente	526,194
4. Forestales	0	4. Forestales	0
5. Pastos	306,885	5. Pastos	291,033
6. Otros Usos	39,286	6. Otros Usos	34,899
7. Bosque Natural	860,668	7. Bosque Natural	853,156
8. No Agropecuario	7,761	8. No Agropecuario	0,460
9. Nubes	18,140	9. Nubes	19,823
10. Cuerpos de Agua	0	10. Cuerpos de Agua	0
11. Suelo en preparación	0	11. Suelo en preparación	0
12. Banco arena	6,411	12. Banco arena	7,041
13. Vegetación Secundaria	333,586	13. Vegetación Secundaria	325,606
14. Sombras	0	14. Sombras	0
Total	2095,416	Total	2095,416

Script V03-03



Script V03-01



- Cobertura del Suelo**
- Clase**
- Cultivo Sin Clasificar
 - Transitorio
 - Permanente
 - Forestales
 - Pastos
 - Otros Usos
 - Bosque Natural
 - No Agropecuario
 - Nubes
 - Cuerpos de Agua
 - Suelo en preparación
 - Banco de Arena
 - Vegetación Secundaria
 - Sombras

Ilustración 13. Resultado general de la clasificación de Tununguá

Versión 03-03

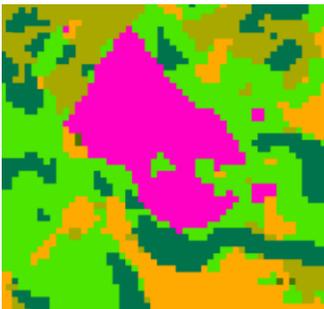


Imagen de mejor resolución



Versión 03-01

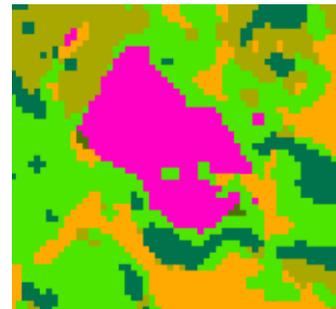


Ilustración 14. Resultados detallados de la clasificación de Tununguá

En la Ilustración 14, se observa como el lote en descaso, identificado en color fucsia, tiene una mejor definición de los bordes en la clasificación realizada por el Script versión 03-03. Igualmente, se observa que los bosques, en color verde oscuro, son más consistentes con la imagen de mejor resolución, mientras que en la versión 03-01 algunos fueron excluidos.



Ilustración 15. Resultados detallados de la clasificación de Tununguá

En la Ilustración 15, se observa de forma más detallada las diferencias entre las dos clasificaciones. En particular, en las coberturas de bosque natural, representadas en verde oscuro, vegetación secundaria, en verde viche, y pastos, en amarillo.

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01
<https://code.earthengine.google.com/bf2bc52b729bcd81c2f9bcdb81af3e4>
- V03-03
<https://code.earthengine.google.com/975c9685b6f62e9d0df29da874677957>

4.3.1.3 Pruebas realizadas en Pajarito

- Pruebas para la zona rural

- Producto del proceso de remuestreo



Ilustración 16. Proceso de remuestreo de la imagen de Pajarito Rural - Combinación RGB

- Muestras de entrenamiento

Las muestras de entrenamiento se capturaron tomando como base la interpretación realizada por un experto del grupo de MMRA, en la que no se identificó ningún tipo de cultivo. La Tabla 16 indica el número de muestras usadas por cada cobertura.

Tabla 16. Muestras de Entrenamiento para Pajarito – zona rural

Cobertura	Número de muestras	Cobertura	Número de muestras
1. Cultivo sin clasificar	0	8. No Agropecuario	13
2. Transitorio	0	9. Nubes	0
3. Permanente	0	10. Cuerpos de Agua	37
4. Forestales	0	11. Suelo en preparación	0
5. Pastos	84	12. Banco arena	0
6. Otros Usos	0	13. Vegetación Secundaria	0
7. Bosque Natural	33	14. Sombras	0

Tabla 17. Índices de calidad de la clasificación de Pajarito - zona rural

Categoría	Script V03-01				Script V03-03			
	Productor	Error de Omisión	Consumidor	Error de comisión	Productor	Error de omisión	Consumidor	Error de comisión
Cultivo sin	0	0	0	0	0	0	0	0
Transitorio	0	0	0	0	0	0	0	0
Permanente	0	0	0	0	0	0	0	0
Forestales	0	0	0	0	0	0	0	0
Pastos	0.9	10%	0,9574	4.26%	1	0%	0,9615	3.85%
Otros Usos	0	0	0	0	0	0	0	0
Bosque Natural	0.9921	0.79%	0.9843	1.57%	0.9921	0.79%	1	0%
No Agropecuario	1	0%	0.8	20%	1	0%	1	0%
Nubes	0	0	0	0	0	0	0	0
Cuerpos de Agua	0.96	4%	1	0%	0.96	4%	1	0%
Suelo en preparación	0	0	0	0	0	0	0	0
Banco arena	0	0	0	0	0	0	0	0
Vegetación Secundaria	0	0	0	0	0	0	0	0
Sombras	0	0	0	0	0	0	0	0
Precisión General	0,9673				0,9907			
Indicador Kappa	0,9433				0,9838			

- Resultados de la clasificación

La Tabla 18, muestra la comparación de las áreas obtenidas de las clasificaciones a partir de los dos scripts, donde se puede observar que, en el script V03-03, el área de las coberturas de pastos, no agropecuario y cuerpos de agua es menor, y el área del bosque natural es mayor. El área total calculada con el script V03-01 es mayor que el área total calculada con el script V03-03, obteniendo una diferencia de 0.18 Has.

Tabla 18. Áreas por cobertura de Pajarito - zona rural

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
1. Cultivo sin	0	1. Cultivo sin	0
2. Transitorio	0	2. Transitorio	0
3. Permanente	0	3. Permanente	0
4. Forestales	0	4. Forestales	0
5. Pastos	125,61	5. Pastos	123,86
6. Otros Usos	0	6. Otros Usos	0
7. Bosque Natural	157,27	7. Bosque Natural	160,13
8. No Agropecuario	13,5	8. No Agropecuario	12,23
9. Nubes	0	9. Nubes	0
10. Cuerpos de Agua	14,44	10. Cuerpos de Agua	14,42
11. Suelo en preparación	0	11. Suelo en preparación	0
12. Banco arena	0	12. Banco arena	0
13. Vegetación Secundaria	0	13. Vegetación Secundaria	0
14. Sombras	0	14. Sombras	0
Total	310,82	Total	310,64

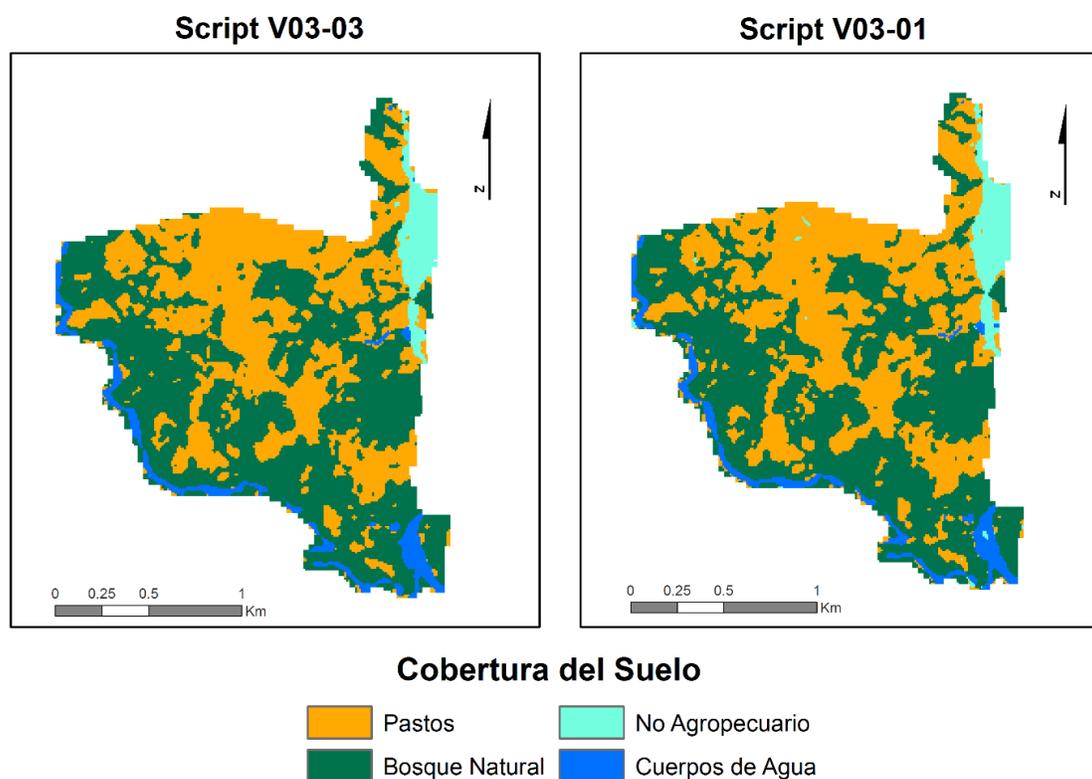


Ilustración 17. Resultado general de la clasificación de Pajarito – zona rural



Ilustración 18. Resultados detallados de la clasificación de Pajarito Rural

En la Ilustración 18. Resultados detallados de la clasificación de Pajarito Rural se observa como los cuerpos de agua, identificados en color azul, tienen mejor consistencia en la clasificación realizada por el Script V03-03, ya que en la clasificación realizada por el Script V03-01, para esta misma cobertura, se evidencia cómo se ve interferida por la cobertura de no agropecuario, la que a su vez tiende a confundirse un poco con la cobertura de pastos. Este resultado visual representa a su vez, los valores obtenidos en los índices de calidad, Tabla 17, para la cobertura no agropecuario, ya que en el script V03-01 presenta un error de comisión del 20% y un error del 0% para la otra versión.

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01

<https://code.earthengine.google.com/b9c8542922d64e96c4bf4275e75a960e>

- V03-03

<https://code.earthengine.google.com/66071f24b477de3fdd5b444b63ddc328>

- Pruebas para la zona urbana

- Producto del proceso de remuestreo



Ilustración 19. Proceso de remuestreo de la imagen de Pajarito Urbano - Combinación RGB

- Muestras de entrenamiento

Las muestras de entrenamiento se capturaron tomando como base la interpretación realizada por un experto del grupo de MMRA, en la que no se identificó ningún tipo de cultivo. La Tabla 19, indica el número de muestras usadas por cada cobertura.

Tabla 19. Muestras de Entrenamiento para Pajarito – zona urbana

Cobertura	Número de muestras	Cobertura	Número de muestras
1. Cultivo sin clasificar	0	8. No Agropecuario	33
2. Transitorio	0	9. Nubes	0
3. Permanente	0	10. Cuerpos de Agua	37
4. Forestales	0	11. Suelo en preparación	7
5. Pastos	50	12. Banco arena	0
6. Otros Usos	74	13. Vegetación Secundaria	0
7. Bosque Natural	32	14. Sombras	0

Tabla 20. Índices de calidad de la clasificación de Pajarito - zona urbana

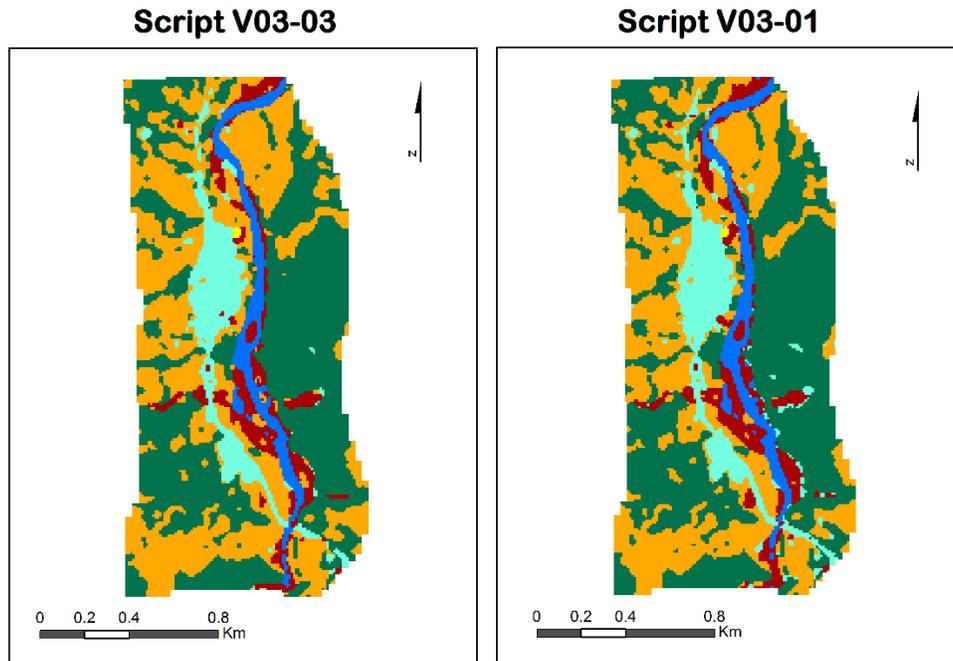
Categoría	Script V03-01				Script V03-03			
	Productor	Error de Omisión	Consumidor	Error de comisión	Productor	Error de omisión	Consumidor	Error de comisión
Cultivo sin	0	0	0	0	0	0	0	0
Transitorio	0	0	0	0	0	0	0	0
Permanente	0	0	0	0	0	0	0	0
Forestales	0	0	0	0	0	0	0	0
Pastos	0.6667	33.33%	0.8571	14.29%	1	0%	0.9	10%
Otros Usos	0.7	30%	0.875	12.5%	0.65	35%	0.8125	18.75%
Bosque Natural	1	0%	1	0%	1	0%	1	0%
No Agropecuario	1	0%	0.7727	22.73%	0.9412	5.88%	0.8421	15.79%
Nubes	0	0	0	0	0	0	0	0
Cuerpos de Agua	0.9412	5.88%	0.9412	5.88%	0.9118	8.82%	0.9688	3.12%
Suelo en preparación	0	0	0	0	0	0	0	0
Banco arena	0	0	0	0	0	0	0	0
Vegetación Secundaria	0	0	0	0	0	0	0	0
Sombras	0	0	0	0	0	0	0	0
Precisión General					0,9405			
Indicador Kappa					0,9044			

- Resultados de la clasificación

La Tabla 21, muestra la comparación de las áreas obtenidas de las clasificaciones a partir de los dos scripts, donde se puede observar que, en el script V03-03, el área de las coberturas de pastos y suelo en preparación es mayor, y el área de las demás coberturas es menor con respecto al script V03-01. El área total calculada con el script V03-01 es mayor que el área total calculada con el script V03-03, obteniendo una diferencia de 0.64 Has.

Tabla 21. Áreas por cobertura de Pajarito - zona urbana

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
1. Cultivo sin	0	1. Cultivo sin	0
2. Transitorio	0	2. Transitorio	0
3. Permanente	0	3. Permanente	0
4. Forestales	0	4. Forestales	0
5. Pastos	74,75	5. Pastos	78,43
6. Otros Usos	14,92	6. Otros Usos	13,73
7. Bosque Natural	95,03	7. Bosque Natural	94,19
8. No Agropecuario	20,87	8. No Agropecuario	18,6
9. Nubes	0	9. Nubes	0
10. Cuerpos de Agua	11,58	10. Cuerpos de Agua	11,53
11. Suelo en preparación	0,15	11. Suelo en preparación	0,18
12. Banco arena	0	12. Banco arena	0
13. Vegetación Secundaria	0	13. Vegetación Secundaria	0
14. Sombras	0	14. Sombras	0
Total	217,30	Total	216,66



Cobertura del Suelo

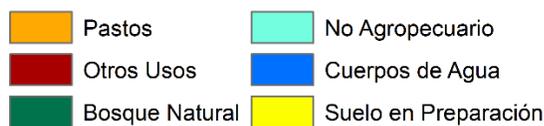


Ilustración 20. Resultado general de la clasificación de Pajarito – zona urbana

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01
<https://code.earthengine.google.com/984ba0e72e22d8c8b07e8c31c7dc5325>
- V03-03
<https://code.earthengine.google.com/f5c37bc628660431aaa8cfbea0e41d4e>

4.3.1.4 Pruebas realizadas en Cota

- Producto del proceso de remuestreo

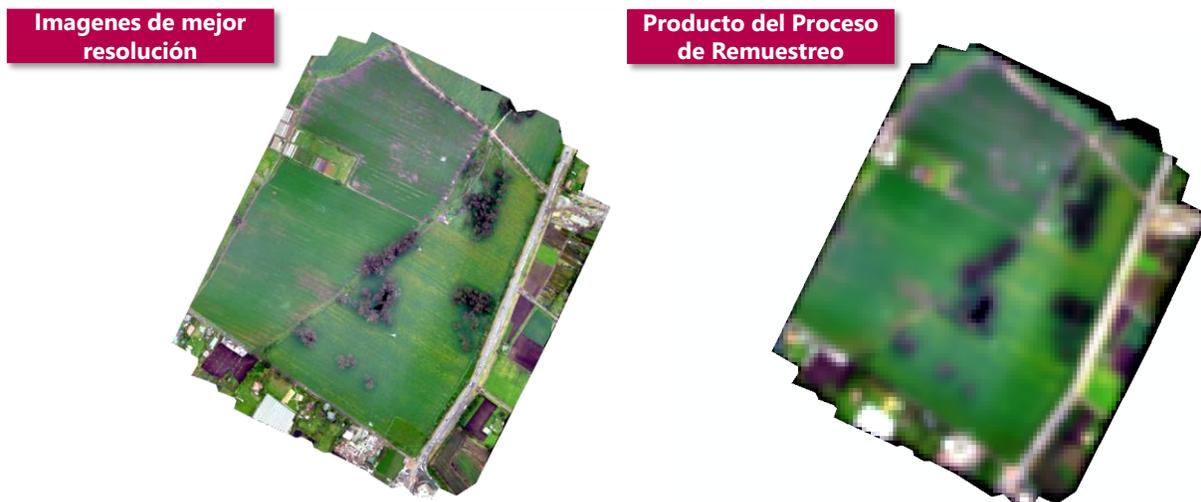


Ilustración 21. Proceso de remuestreo de la imagen de Cota - Combinación RGB

- Muestras de entrenamiento

La imagen de Cota contiene un área muy pequeña en la cual solo se identificó la presencia de algunos cultivos, por esta razón, las pruebas se realizaron con el fin de evaluar únicamente la capacidad de discriminación de los scripts para dichos cultivos de forma desagregada. Además, como consecuencia de que las imágenes satelitales de Sentinel-2 corresponden a fechas muy lejanas con respecto a la fecha de captura de la imagen del RPAS, las coberturas presentes en ambas imágenes muestran cambios en su estructura con el tiempo, como se puede evidenciar en la Ilustración 22, por lo tanto, fue necesario tomar muestras de entrenamiento diferentes para cada versión del script, las cuales se capturaron tomando como base la interpretación realizada por un experto del grupo de MMRA. La Tabla 22 y la Tabla 23 indican el número de muestras usadas por cada cobertura en cada versión del script.



Ilustración 22. Comparación de coberturas en imagen Sentinel-2 e imagen de RPAS para Cota

Tabla 22. Muestras de entrenamiento para Cota con Script V03-01

Cobertura	Número de muestras
1. Pastos	37
2. Maíz	98
3. Otros Usos	44
4. Hortalizas	46
5. Barbecho	54

Tabla 23. Muestras de entrenamiento para Cota con Script V03-03

Cobertura	Número de muestras
1. Pastos	72
2. Maíz	129
3. Otros Usos	44
4. Hortalizas	58
5. Barbecho	52

Tabla 24. Índices de calidad de la clasificación de Cota

Categoría	Script V03-01				Script V03-03			
	Productor	Error de Omisión	Consumidor	Error de comisión	Productor	Error de omisión	Consumidor	Error de comisión
Pastos	0,875	12.5%	1	0%	0,6923	30.77%	0,9	10%
Maíz	0,9555	4.45%	0,7166	28.34%	0,998	0.2%	0,9592	4.08%
Otros Usos	1	0%	0,1645	83.55%	0,9444	5.56%	1	0%
Hortalizas	0,1111	88.89%	0,0149	98.51%	0,2222	77.78%	1	0%
Barbecho	0,6375	36.25%	0,9661	3.39%	0,8275	17.35%	0,923	7.7%
Precisión General	0,7165				0,9581			
Indicador Kappa	0,51				0,8006			

- Resultados de la clasificación

La Tabla 25, muestra las áreas obtenidas de las clasificaciones a partir de los dos scripts, donde se puede observar que, en el script V03-03, el área de la cobertura de maíz es mucho mayor, mientras que, en el script V03-01 es mucho mayor el área de la cobertura de barbecho.

Tabla 25. Áreas por cobertura de Cota

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
Pastos	1.68	Pastos	2.84
Maíz	38.25	Maíz	51.69
Otros Usos	5.79	Otros Usos	5.87
Hortalizas	1.53	Hortalizas	1.24
Barbecho	17.31	Barbecho	2.92
Total	64.56	Total	64.56

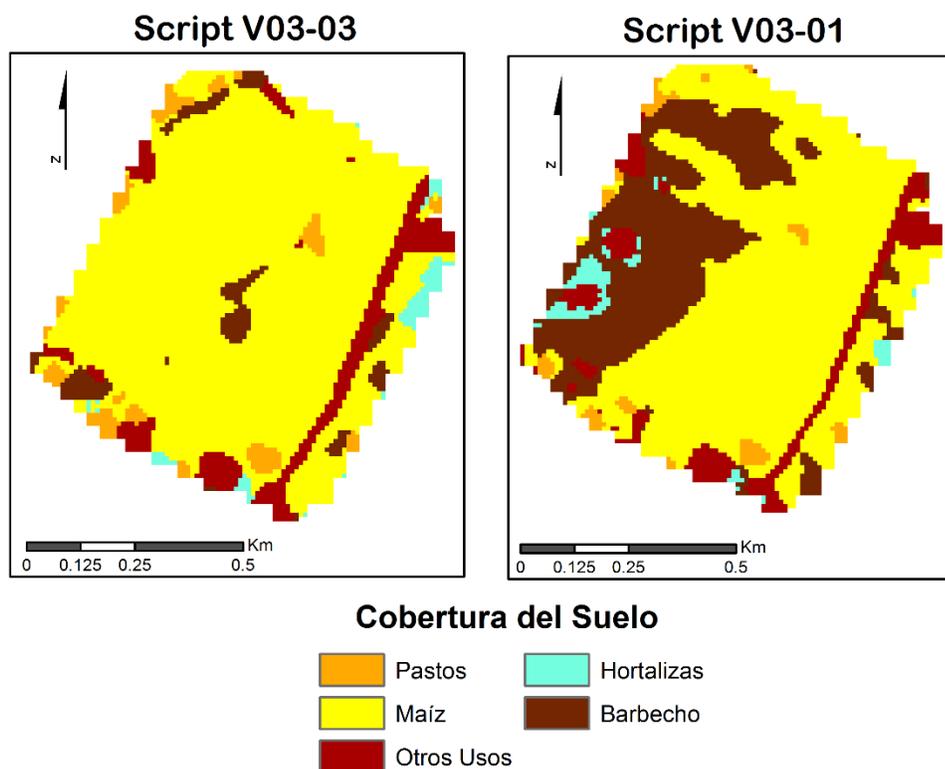


Ilustración 23. Resultado general de la clasificación de Cota

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01
<https://code.earthengine.google.com/2d119630175a78be20e0637bfe2607be>
- V03-03
<https://code.earthengine.google.com/ffd5fe60e51bf050eda845908a5f793d>

Bosque Natural	0,9	10%	1	0%	0,9	10%	1	0%
No Agropecuario	1	0%	1	0%	1	0%	1	0%
Nubes	0	0	0	0	0	0	0	0
Cuerpos de Agua	0	0	0	0	0	0	0	0
Suelo en preparación	1	0%	0,75	25%	1	0%	0,75	25%
Banco arena	0	0	0	0	0	0	0	0
Vegetación Secundaria	0	0	0	0	0	0	0	0
Sombras	0	0	0	0	0	0	0	0
Precisión General				0,931				0,931
Indicador Kappa				0,906				0,903

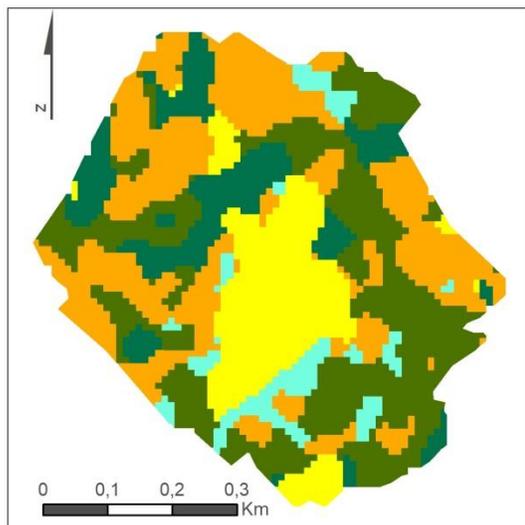
- Resultados de la clasificación

La Tabla 28, muestra las áreas obtenidas de las clasificaciones a partir de los dos scripts, donde se puede observar que, en el script V03-03, aumento el área para las coberturas de transitorio y no agropecuario, mientras que, se presenta una disminución en las áreas para pastos, bosque natural y suelo en preparación con respecto a la clasificación realizada con la V03-01.

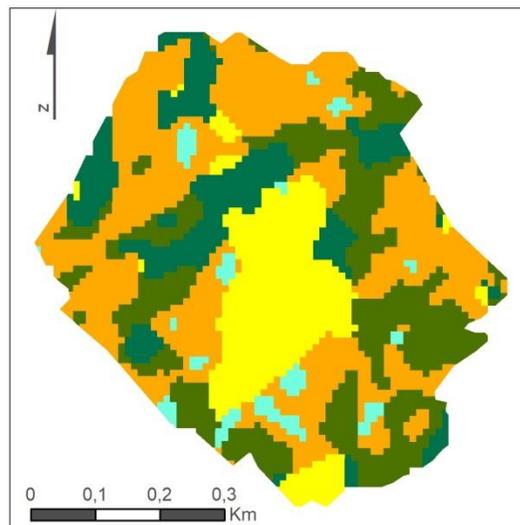
Tabla 28. Áreas por cobertura de Cómbita

Versión 03-01		Versión 03-03	
Cobertura	Área (ha)	Cobertura	Área (ha)
1. Cultivo sin	0	1. Cultivo sin	0
2. Transitorio	8,915	2. Transitorio	10,801
3. Permanente	0	3. Permanente	0
4. Forestales	0	4. Forestales	0
5. Pastos	13,592	5. Pastos	11,033
6. Otros Usos	0	6. Otros Usos	0
7. Bosque Natural	4,854	7. Bosque Natural	4,684
8. No Agropecuario	1,229	8. No Agropecuario	2,106
9. Nubes	0	9. Nubes	0
10. Cuerpos de Agua	0	10. Cuerpos de Agua	0
11. Suelo en preparación	5,881	11. Suelo en preparación	5,8
12. Banco arena	0	12. Banco arena	0
13. Vegetación Secundaria	0	13. Vegetación Secundaria	0
14. Sombras	0	14. Sombras	0
Total	34,473	Total	34,473

Script V03-03



Script V03-01



Cobertura del Suelo

 Cultivo Sin Clasificar	 No Agropecuario
 Transitorio	 Nubes
 Permanente	 Cuerpos de Agua
 Forestales	 Suelo en preparación
 Pastos	 Banco de Arena
 Otros Usos	 Vegetación Secundaria
 Bosque Natural	 Sombras

Ilustración 25. Resultado general de la clasificación de Cómbita

- Evidencias

Las evidencias de los resultados anteriormente mencionados se encuentran en los siguientes links

- V03-01

<https://code.earthengine.google.com/c65f2ad0c4e1e85dc48d588ad6788fb8>

- V03-03

<https://code.earthengine.google.com/8fe7fa27667e1bf47c962e9f7717c64a>

5. ANÁLISIS DE RESULTADOS

El análisis de los resultados está enfocada en los aspectos más relevantes buscados con la ejecución del proyecto, los cuales son a: (i) comportamiento de las coberturas, (ii) comparación de los resultados de la clasificación solo sentinel-2 y el flujo metodológico propuesto que integra la información del RPAS y por último (iii) el grado de mejora que evidencia los resultados en comparación a los obtenidos con los métodos de fotointerpretación manual. Por lo anterior fueron realizadas las siguientes tablas, las cuáles respectivamente muestran la información relevante de los aspectos mencionados previamente.

Tabla 29. Síntesis de errores de comisión

	Error de Comisión									
	Script V03-01					Script V03-03				
	Salento	Tununguá	Pajarito Rural	Pajarito Urbano	Cómbita	Salento	Tununguá	Pajarito Rural	Pajarito Urbano	Cómbita
1. Cultivo Sin	-	2,5%	-	-	-	-	3,7%	-	-	-
2. Transitorio	47,1%	18,6%	-	-	20%	23,1%	10,3%	-	-	0%
3. Permanente	10,1%	31,2%	-	-	-	9,5%	29,9%	-	-	-
4. Forestales	-	-	-	-	-	-	-	-	-	-
5. Pastos	2,0%	21,1%	4.26%	14.29%	0%	4,0%	21,6%	3.85%	10%	9%
6. Otros Usos	-	0%	-	12.50%	-	-	0%	-	18.75%	-
7. Bosque Natural	1,2%	4,5%	1.57%	0%	0%	1,0%	3,8%	0%	0%	0%
8. No Agropecuario	0%	88,9%	20%	22.73%	0%	4,3%	25,0%	0%	15.79%	0%
9. Nubes	-	1,0%	-	-	-	-	-	-	-	-
10. Cuerpos de Agua	6,2%	-	0%	5.88%	-	13,9%	-	0%	3.12%	-
11. Suelo en Preparación	-	-	-	-	25%	-	-	-	-	25%
12. Banco de Arena	33,3%	0%	-	-	-	20,0%	0%	-	-	-
13. Vegetación Secundaria	-	30,2%	-	-	-	-	28,4%	-	-	-
14. Sombras	-	-	-	-	-	-	-	-	-	-

Tabla 30. Síntesis de los errores de omisión

	Error de Omisión									
	Script V03-01					Script V03-03				
	Salento	Tununguá	Pajarito Rural	Pajarito Urbano	Cómbita	Salento	Tununguá	Pajarito Rural	Pajarito Urbano	Cómbita
1. Cultivo Sin	-	3,7%	-	-	-	-	3,7%	-	-	-
2. Transitorio	9,8%	30,4%	-	-	0%	2,4%	24,6%	-	-	20%
3. Permanente	8,9%	26,4%	-	-	-	6,1%	25,3%	-	-	-
4. Forestales	-	-	-	-	-	-	-	-	-	-
5. Pastos	27,9%	19,6%	10%	33.33%	9%	29,4%	17,8%	0%	0%	0%
6. Otros Usos	-	3,3%	-	30%	-	-	3,3%	-	35%	-
7. Bosque Natural	1,6%	4,0%	0.79%	0%	10%	0,8%	0%	0.79%	0%	10%
8. No Agropecuario	46,9%	33,3%	0%	0%	0%	31,2%	0%	0%	5.88%	0%
9. Nubes	-	23,4%	-	-	-	-	14,5%	-	-	-
10. Cuerpos de Agua	9,1%	-	4%	5.88%	-	6,1%	-	4%	8.82%	-
11. Suelo en Preparación	-	-	-	-	0%	-	-	-	-	0%

12. Banco de Arena	11,1%	0%	-	-	-	11,1%	0%	-	-	-
13. Vegetación Secundaria	-	30,2%	-	-	-	-	31,6%	-	-	-
14. Sombras	-	-	-	-	-	-	-	-	-	-

Los trabajos previos disponibles, (Dash et al., 2018; Zhao et al., 2019), que tratan la clasificación supervisada con *Random Forest* en imágenes Formosat 2 y sentinel-2, buscaron la obtención de coberturas más generales y no en específico cultivos, por lo anterior, la primera parte de la discusión se centra en analizar solamente los resultados con el conocimiento que se posee de las zonas de estudio.

A partir de los resultados de las clasificaciones realizadas en este estudio, se determinan diferentes comportamientos que tienen algunas coberturas presentes en las imágenes y cómo estos se ven reflejados una vez se ha asignado cada píxel a una categoría. Para las zonas de Salento (clasificación de cultivos agregados), y Tununguá, que cuentan con la presencia de cultivos transitorios y permanentes y, por lo tanto, los resultados son comparables entre sí, se tiene que el error de comisión, consignado en la Tabla 29, es significativamente mayor en la clasificación realizada con el script V03-01, con excepción de la cobertura de pastos para la zona de Salento donde el error es mayor un 2% en el script V03-03, resultado que se adjudica al número de muestras tomadas para dicha cobertura que, en el caso de Salento, fueron 73, correspondientes solo al 14% del total de las muestras tomadas, y en Tununguá, fueron 81, correspondientes a un 19% del total de las muestras, situación que se presenta de forma similar en el error de omisión, consignado en la Tabla 30.

Las coberturas de bosque natural, usos no agropecuarios y bancos de arena, coberturas presentes en ambas imágenes, si presentan diferencias significativas para estos valores, siendo el error mucho más alto, tanto de omisión como de comisión, en el script V03-01.

Los resultados anteriores permiten definir que el uso de la imagen de mejor resolución espacial, script v03-03, es un insumo que realmente permite obtener mejores resultados para zonas amplias con coberturas heterogéneas.

En las imágenes capturadas de la zona de Pajarito, divididas en uso urbano y rural, se tiene que, para la cobertura de pastos, el error de omisión se presenta únicamente en el script V03-01, siendo además valores considerables; el error de comisión se presenta en ambas versiones del script, siendo ya un error aludido a la delimitación hecha por el intérprete, sin embargo, este sigue siendo mayor para la versión anteriormente nombrada. La cobertura de otros usos, presente únicamente en la zona urbana, en ambos errores generó valores mayores para la clasificación V03-03, ya que al llevarse a cabo con esta versión tendió a confundirse un poco más con la cobertura de suelo en preparación, sin embargo, esto se confiere a que, como las muestras de entrenamiento tomadas para el suelo en preparación fueron solo 7, en ninguna de la versiones tomó alguna de estas muestras como datos de validación, por lo tanto, el script no se dio la oportunidad de generar datos que corroboraran la clasificación de para este uso. Las demás coberturas identificadas, bosque natural, usos no agropecuarios, y cuerpos de agua, presentan valores parecidos de error de omisión en ambas versiones, por lo que, los scripts generan resultados similares cuando se trata de coberturas que no tienden a variar su estructura significativamente con el paso del tiempo. El error de comisión para

estas mismas coberturas si fue mayor en V03-01, por lo que se puede decir que el script V03-03 tiende a dejar de lado los errores cometidos por el intérprete.

La zona de Cota no es tenida en cuenta para llevar a cabo análisis, como consecuencia de que las imágenes de ambos insumos corresponden a fechas de captura muy diferentes y las coberturas presentan cambios en su estructura no comparables entre sí. La zona de Cóbbita, al corresponder a un área tan pequeña, no permite tomar un número de muestras suficientes tal que el algoritmo pueda entrenarse de forma idónea, por lo que los resultados no generan valores reveladores, más allá de una distinción visual de las coberturas.

Tabla 31. Síntesis de exactitud temática

		Script V03-01					
		Salento	Tununguá	Pajarito Rural	Pajarito Urbano	Cota	Cóbbita
Precisión General		0,922	0,8	0,9673	0,9405	0,7165	0,931
Indicador Kappa		0,878	0,761	0,9433	0,9044	0,51	0,906
		Script V03-03					
		Salento	Tununguá	Pajarito Rural	Pajarito Urbano	Cota	Cóbbita
Precisión General		0,942	0,823	0,9907	0,9405	0,9581	0,931
Indicador Kappa		0,909	0,787	0,9838	0,9048	0,8006	0,903

La Tabla 31, indica que en un promedio de 3% los indicadores de exactitud temática mejoraron en comparación de los scripts versión 03-01 (Solo sentinel-2 y pansharpening) y 03-03 (Sentinel-2 integrado con RPAS), aunque la mejora puede parecer no significativa, uno de los aspectos que se debe considerar como se mencionó previamente, es referente al mejoramiento de la delimitación de cada una de las coberturas en todas las zonas de estudio visto de forma visual y también evidenciados en los errores de comisión y omisión de las coberturas objetivo, inclusive fue notoria esta característica en las de menor área como lo son la zona de Cota y Combita.

Se debe especificar que la fusión de la imagen sentinel-2 y RPAS tiene el objetivo de aumentar la resolución espectral del insumo objetivo (mosaico sentinel-2), siendo el mismo que busco Zhao et al., (2019) quienes realizaron una investigación similar a la presentada, pero su propósito fue fortalecer la resolución espectral a resoluciones espaciales cercanas a las de las imágenes RPAS con la información multispectral de las imágenes Sentinel-2 buscando que fuera efectiva esta característica en los resultados de la clasificación supervisada a nivel de cultivos desagregados, siendo este mismo el proceso decisivo de la investigación presentada.

No obstante, a partir de los resultados del trabajo de Zhao et al., (2019) se evidencia que lo que para ellos fue un inconveniente (aplicar el método en resoluciones espaciales alejadas de la original del RPAS), fue una fortaleza en esta investigación, debido a que, de acuerdo con los resultados de la mayoría de los experimentos realizados con cultivos desagregados arrojaron índices de calidad cercanos al 90% mientras que en el estudio referenciado fue en promedio del 75%, además de demostrar efectividad en áreas de estudios amplias (Salento y Tunungua) característica no observada en las investigaciones previas. Por otro lado, se debe mencionar que como debilidad de la investigación realizada y de las mencionadas radica que, con las condiciones dispuestas, no es

posible extrapolar la información precisa que contiene la imagen de mejor resolución espacial a otras zonas geográficas tomadas solamente por la imagen satelital analizada.

Por otra parte, los resultados obtenidos a partir del proceso de actualización con la versión 03-03 fueron comparados con el proceso de fotointerpretación realizado en el año 2019 para realizar el proceso de actualización del Marco Maestro Rural Agropecuario. El Marco MMRA se consolidó en el año 2017 usando como base el Censo Nacional Agropecuario. Posteriormente, en el año 2019 se actualizó los porcentajes de las coberturas para los departamentos de Boyacá, Nariño, Norte de Santander y algunos municipios de Cundinamarca fotointerpretando imágenes de Google Earth del año 2017, En concreto, los expertos determinaban visualmente el porcentaje de una cobertura y lo consignaban en una tabla asociada a cada conglomerado. En la Tabla 32 y en la Tabla 33 se comparan los porcentajes obtenidos a partir del proceso de fotointerpretación y los porcentajes obtenidos a partir del proceso de clasificación con el script V03-03 para los conglomerados totalmente contenidos en el área de estudio del municipio de Tununguá y Pajarito.

Tabla 32. Comparación de resultados del método de clasificación con el método de fotointerpretación en Tununguá

Conglomerado	área(ha)	Método	CULTIVOS SIN CLASIFICAR	TRANSITORIOS	PERMANENTES	FORESTALES	PASTOS	OTROS USOS	BOSQUES	NO AGRO PECUARIOS
158320005000	51,78	Clasificación	0,09	0	0,26	0	0,13	0,27	0,25	0
		Fotointerpretación	0	0	0,3	0	0,6	0	0,1	0
		Diferencia	9,0%	0,0%	-4,0%	0,0%	-47,0%	27,0%	15,0%	0,0%
158320006000	46,48	Clasificación	0	0	0,24	0	0,05	0,16	0,54	0
		Fotointerpretación	0	0	0,3	0	0,3	0	0,4	0
		Diferencia	0,0%	0,0%	-6,0%	0,0%	-25,0%	16,0%	14,0%	0,0%
158320007000	188,36	Clasificación	0	0	0,33	0	0,15	0,2	0,32	0
		Fotointerpretación	0	0	0,35	0	0,5	0,05	0,1	0
		Diferencia	0,0%	0,0%	-2,0%	0,0%	-35,0%	15,0%	22,0%	0,0%
158320011000	51,17	Clasificación	0,02	0,02	0,5	0	0,2	0,14	0,12	0
		Fotointerpretación	0	0	0,15	0	0,65	0	0,2	0
		Diferencia	2,0%	2,0%	35,0%	0,0%	-45,0%	14,0%	-8,0%	0,0%
158320013000	210,36	Clasificación	0,01	0	0,17	0	0,11	0,18	0,53	0
		Fotointerpretación	0	0	0,25	0	0,5	0	0,25	0
		Diferencia	1,0%	0,0%	-8,0%	0,0%	-39,0%	18,0%	28,0%	0,0%
158320020000	47,19	Clasificación	0,01	0	0,09	0	0,16	0,29	0,45	0
		Fotointerpretación	0	0	0,1	0	0,7	0	0,2	0
		Diferencia	1,0%	0,0%	-1,0%	0,0%	-54,0%	29,0%	25,0%	0,0%
158320023000	45,85	Clasificación	0	0,04	0,57	0	0,06	0,17	0,17	0
		Fotointerpretación	0	0,05	0,45	0	0,35	0	0,15	0
		Diferencia	0,0%	-1,0%	12,0%	0,0%	-29,0%	17,0%	2,0%	0,0%

158320028000	50,86	Clasificación	0	0,02	0,38	0	0,18	0,12	0,29	0
		Fotointerpretación	0	0	0,6	0	0,25	0	0,15	0
		Diferencia	0,0%	2,0%	-22,0%	0,0%	-7,0%	12,0%	14,0%	0,0%
158320032000	202,72	Clasificación	0,01	0	0,25	0	0,23	0,16	0,35	0
		Fotointerpretación	0	0	0,4	0	0,4	0	0,2	0
		Diferencia	1,0%	0,0%	-15,0%	0,0%	-17,0%	16,0%	15,0%	0,0%
158320035000	52,69	Clasificación	0	0	0,23	0	0,18	0,11	0,48	0
		Fotointerpretación	0	0	0,1	0	0,65	0	0,25	0
		Diferencia	0,0%	0,0%	13,0%	0,0%	-47,0%	11,0%	23,0%	0,0%
158320037000	24,69	Clasificación	0,03	0,02	0,63	0	0,06	0,02	0,25	0
		Fotointerpretación	0	0	0,25	0	0,5	0,05	0,2	0
		Diferencia	3,0%	2,0%	38,0%	0,0%	-44,0%	-3,0%	5,0%	0,0%
158320038000	109,69	Clasificación	0,05	0	0,04	0	0,12	0,09	0,7	0
		Fotointerpretación	0	0	0,05	0	0,4	0,05	0,5	0
		Diferencia	5,0%	0,0%	-1,0%	0,0%	-28,0%	4,0%	20,0%	0,0%
158320040000	43,05	Clasificación	0	0	0,23	0	0,09	0,14	0,54	0
		Fotointerpretación	0	0	0,1	0	0,55	0,05	0,3	0
		Diferencia	0,0%	0,0%	13,0%	0,0%	-46,0%	9,0%	24,0%	0,0%

Tabla 33. Comparación de resultados del método de clasificación con el método de fotointerpretación en Pajarito

Conglomerado	área(ha)	Método	CULTIVOS SIN CLASIFICAR	TRANSITORIOS	PERMANENTES	FORESTALES	PASTOS	OTROS USOS	BOSQUES	NO AGRO PECUARIOS	No identificado
158320005000	283,75	Clasificación	0	0	0	0	0,35	0,06	0,53	0,02	0,04
		Fotointerpretación	0	0	0	0	0,8	0	0,2	0	0
		Diferencia	0,0%	0,0%	0,0%	0,0%	-44,8%	6,0%	33,3%	1,9%	3,6%

A partir de la comparación realizada en Tununguá y Pajarito entre los dos métodos, se puede identificar que en las clasificaciones realizadas con el script V03-03, la cobertura pastos reporto un porcentaje de área menor en todos los conglomerados, mientras que las coberturas otros usos y bosques tiene un porcentaje área mayor. Tal vez esta situación se deba a que en la clasificación realizada con el script tenga dificultades para identificar los pastos, pero si se observan los índices de calidad estos indican que el porcentaje de omisión para esta clase es del 17,8% en Tununguá y del 0% en Pajarito, lo cual no explica totalmente la diferencia del resultado, Por tanto, se puede sugerir que la diferencia se deba a la subjetividad del interprete, quien posiblemente tienda a favorecer en la clase pastos y prestar menos interés en las clase bosques y/o otros usos.

En el caso del área de estudio de Tununguá, la clase cultivos permanente tiene un comportamiento dinámico. De acuerdo con la clasificación realizada con el script V03-03 en algunos conglomerados se presentó un porcentaje de área mayor, mientras que en otros se presentó un reporto un porcentaje menor con respecto a los porcentajes reportados en el proceso de fotointerpretación. Considerando

que en esta clase se presentó un error de omisión del 25,3 % y un error de comisión del 29,9% es coherente el anterior comportamiento, aunque esto no descarta que la diferencia de deba a la subjetividad del interprete.

Por otra parte, se puede identificar una diferencia entre el número de clases identificadas en cada uno de los métodos. En el caso del método de clasificación con GEE, en doce conglomerados de los catorce conglomerados comparados reportan un número mayor de clases con respecto al proceso de foto interpretación, en los otros dos conglomerados se reportan el mismo número de clases. Se infiere que el proceso de clasificación semiautomatizado en GEE permite la discriminación de clases que en el proceso de fotointerpretación son omitidas, dado que en el proceso de fotointerpretación el análisis se hace sobre imágenes con una combinación en color verdadero, mientras que la clasificación en GEE incorpora un mayor número de bandas, el modelo digital del terreno e índices.

6. EVALUACIÓN Y CUMPLIMIENTO DE LOS OBJETIVOS DE LA PASANTÍA

OBJETIVO	ACTIVIDAD	CUMPLIMIENTO
Revisar y analizar fuentes que permitan establecer el uso de información de RPAS para mejorar la clasificación de coberturas	Recopilar información para su posterior análisis	100%
Contribuir al desarrollo de la metodología y el algoritmo de clasificación automatizada de coberturas del MMRA en la plataforma Google Earth Engine utilizando imágenes disponibles de Sentinel-2.	Probar el script, proponer mejoras, consolidar los pasos a realizar e identificar los actores que intervienen en el proceso	100%
Clasificar las coberturas de las áreas seleccionadas del MMRA mediante imágenes de dron e imágenes de satélite	Clasificar las coberturas de imágenes Sentinel e imágenes de dron a través Google Earth Engine	100%
Contribuir a la generación de un artículo científico en el que se conceptualice la metodología propuesta y los resultados obtenidos	Proporcionar los conocimientos generados a través del uso del script para la clasificación de las coberturas	100%

7. CONCLUSIONES Y RECOMENDACIONES

El proceso de clasificación supervisada de coberturas semiautomatizado en la plataforma Google Earth Engine y el uso de algoritmos de *Machine Learning*, representa una mejora en el proceso de actualización de la información agropecuaria a nivel Nacional, en particular en aspectos como: la reducción de subjetividad, puesto que el intérprete solo interviene capturando las muestras de entrenamiento, los demás subprocesos están parametrizados, garantizando que todas clasificaciones se realicen con los mismos criterios; por otro lado, la metodología propuesta incorpora hasta veintinueve variables en el algoritmo de clasificación, por lo tanto, aumenta la probabilidad de tener una clasificación aceptable y la identificación de clases que por medio de fotointerpretación son omitidas.

Para aquellas zonas que cuentan con presencia de cultivos transitorios y/o permanentes, contar con una imagen adquirida mediante un RPAS supone una enorme ventaja en términos de precisión en la delimitación de las coberturas e identificación de aquellas que se encuentran dispuestas en áreas pequeñas, como fue el caso de Salento. Por otro lado, en las demás zonas de estudio, la imagen de mejor resolución espacial permite definir de forma más precisa los bordes de un cuerpo de agua, de una carretera, las fronteras entre cubiertas con características a simple vista similares, como los bosques y la vegetación secundaria, entre otros, todo esto demostrado en los resultados de las clasificaciones realizadas.

A partir del análisis de los resultados obtenidos de la clasificación de cada una de las zonas propuestas, se establece una consideración relevante a ser tomada en cuenta para el aprovechamiento del script que integra la información del satélite y del RPAS, y hace referencia a la importancia de contar con información con fechas de adquisición cercanas entre ellas, con el fin de que el estado de las coberturas bajo estudio sea comparable y no posean modificaciones en su estructura y varíe de una imagen a otra.

Sin duda, se puede indicar que la integración de imágenes de muy alta resolución espacial con una imagen satelital de alta resolución (siendo en este caso RPAS y Sentinel-2 respectivamente), mediante un filtro de textura gaussiano y remuestreo iterativo, el cual fortalece la resolución espectral, proporciona mejores resultados en la clasificación supervisada con respecto a los obtenidos con un enfoque de solo imágenes satelitales. No obstante, se debe mencionar que este mejoramiento solo está limitado a las zonas geográficas cubiertas completamente por estos dos productos y que puede variar en cierto grado los resultados dependiendo de las características propias de las imágenes de mejor resolución espacial.

Por lo anterior, se recomienda el fortalecimiento de este tipo de investigaciones con el estudio y exploración de métodos estadísticos que permitan la extrapolación de la información precisa de las imágenes de mejor resolución espacial a zonas que solamente posean toma por parte de la imagen satelital. Además de contemplar la inclusión de imágenes de sensores no ópticos en especial para la generación de información geoespacial de las zonas del país que poseen nubosidad constante.

8. BIBLIOGRAFÍA

- Aeronáutica Civil, 1 (2018). [http://www.aerocivil.gov.co/normatividad/Resoluciones TA 2018/RESL. N° 04201 DIC 27 de 2018.pdf](http://www.aerocivil.gov.co/normatividad/Resoluciones_TA_2018/RESL_Nº_04201_DIC_27_de_2018.pdf)
- Albacete, A. (2019). *Tratamiento de imágenes Sentinel-2 con ArcGIS*. <https://storymaps.arcgis.com/stories/8cf7c8024f6e4eb1bbb1d977bda137cf>
- Arango, M., Branch, J., & Botero, V. (2005). Clasificación no supervisada de coberturas vegetales sobre imágenes digitales de sensores remotos: “LANDSAT– ETM+.” *Facultad Nacional de Agricultura Medellín*, 58(1), 2611–2634.
- Ardila, J. P., Espejo, O. J., & Herrera, J. (2005). Validación de una Metodología de Clasificación de Imágenes Satelitales en un entorno Orientado a Objetos. *Ingeniería*, 10(1), 61–69. <https://doi.org/10.14483/23448393.1877>
- Ariputhiran, G., & S. Gandhimathi, @usha. (2013). Feature extraction and classification of high resolution satellite images using GLCM and Back Propagation technique. *International Journal Of Engineering And Computer Science*, 2(2), 525–528. www.ijecs.in
- Ávila Herrera, C. C. (2017). Drones vs. Aeronáutica Civil Licencias para pilotos y su procedimiento. *Novum Jus*, 135–165. <https://doi.org/10.14718/novumjus.2017.11.2.6>
- Balzter, H., Cole, B., Thiel, C., & Schmullius, C. (2015). Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests. *Remote Sensing*, 7(11), 14876–14898. <https://doi.org/10.3390/rs71114876>
- Belgiu, M., & Csillik, O. (2017). Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment*, 204(2018), 509–523. <https://doi.org/10.1016/j.rse.2017.10.005>
- Belgiu, M., & Drăgu, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Bense, T. (2007). Introducción a la percepción remota. *VI Jornadas de Educación En Percepción Remota En El Ámbito Del Mercosur y I Uruguayas de Educación En Percepción Remota*. <http://www.teledet.com.uy/tutorial-imagenes-satelitales/satelites-resolucion-espacial.htm>
- Cabrera, E., Galindo, G., & Vargas, D. M. (2011). *Protocolo de Procesamiento Digital de Imágenes para la Cuantificación de la Deforestación en Colombia, Nivel Nacional Escala Gruesa y Fina*. http://www.ideam.gov.co/documents/13257/13817/Protocolo_para_la_cuantificación_Deforestación_Nivel_Nacional.pdf
- Caseras Saez, I. (2017). *Seguimiento de viñas mediante imágenes multispectrales aéreas y de satélite* [Universidad Pública de Navarra]. <https://pdfs.semanticscholar.org/def5/7d1f16f271995bfb5fdc085165b76d5a3677.pdf>
- Cede, R., Cabrera, E., Vargas, D. M., Galindo, G., García, M. C., Ordoñez, M. F., & Alonso, F. (2011). Introducción a la percepción remota y sus aplicaciones Geológicas. *SIGMUR. SIG y Teledetección En La Universidad de Murcia*, 84(4), 3–4. <https://doi.org/10.4067/S0716-078X2011000400004>
- Chen, W., Xie, X., Wang, J., Pradhan, B., Hong, H., Bui, D. T., Duan, Z., & Ma, J. (2017). A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. *Catena*, 151, 147–160. <https://doi.org/10.1016/j.catena.2016.11.032>
- Chuvieco, E. (2008). *Teledeteccion ambiental*. Grupo Planeta.

- <https://books.google.es/books?id=aKsNXCVCtcQC&printsec=frontcover&dq=teledeteccion+ambiental&hl=es&sa=X&ved=0ahUKEwiE6vLagfbOAhULkRQKHQg2AhMQ6AEIGzAA#v=onepage&q=teledeteccion+ambiental&f=false>
- CORPOICA. (2015). *Ciencia, Tecnología e Innovación en el Sector Agropecuario (Diagnóstico para la Misión para la Transformación del Campo)*. <https://colaboracion.dnp.gov.co/CDT/Agriculturapecuarioforestal+y+pescas/Diagnostico+de+la+Ciencia,+Tecnología+e+Innovación+en+el+Sector+Agropecuario-CORPOICA.pdf>
- DANE. (2016). *Documento Técnico Publicación De Resultados Ena-2016*. https://www.dane.gov.co/files/investigaciones/agropecuario/enda/ena/2016/documento_tecnico_ena_2016.pdf
- DANE. (2019a). *Marco Maestro Rural Y Agropecuaria - Conceptualización Básica*.
- DANE. (2019b). *Metodología General Encuesta Nacional Agropecuaria-ENA*.
- DANE. (2020). *Boletín Técnico Encuesta Nacional Agropecuaria (ENA)*. https://www.dane.gov.co/files/investigaciones/agropecuario/enda/ena/2019/boletin_ena_2019-I.pdf
- Dash, J. P., Pearse, G. D., & Watt, M. S. (2018). UAV multispectral imagery can complement satellite data for monitoring forest health. *Remote Sensing*, *10*(8), 1–22. <https://doi.org/10.3390/rs10081216>
- Díaz García-Cervigón, J. J. (2015). *Estudio de Índices de vegetación a partir de imágenes aéreas tomadas desde UAS/RPAS y aplicaciones de estos a la agricultura de precisión* [Universidad Complutense de Madrid]. https://eprints.ucm.es/31423/1/TFM_Juan_Diaz_Cervignon.pdf
- EMBENTION. (2016). *NM& F300, Nuevo UAV Ala Fija De Alto Rendimiento*. <https://www.interempresas.net/Electronica/Articulos/171969-El-compromiso-dron-Cargas-de-pago-y-la-importancia-de-la-mision.html>
- ESA. (2015). *Sentinel 2*. https://www.esa.int/Space_in_Member_States/Spain/SENTINEL_2
- ESA. (2020). *El programa Copérnico*. https://www.esa.int/Space_in_Member_States/Spain/El_programa_Copernico
- FAO. (2013). *Plan de acción de la estrategia global para el mejoramiento de las estadísticas agropecuarias y rurales*.
- Flood, N. (2013). Seasonal Composite Landsat TM/ETM+ Images Using the Medoid (a Multi-Dimensional Median). *Remote Sensing*, *5*(12), 6481–6500. <https://doi.org/10.3390/rs5126481>
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, *80*(1), 185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Gallo, J. (2002). *Segmentación de Imágenes de Satélite usando Modelos Ocultos de Campos Markovianos*. Centro de Investigación en Matemáticas, A.C.
- García Labrador, M., Évora Brondo, J. A., & Arbelo Pérez, M. (2012). *Satélites de Teledetección para la Gestión del Territorio*.
- González, L. (2018). *Introducción a Machine Learning*.
- HC, T. (2019, July 17). *Pansharpening Sentinel-2 imagery in Google Earth Engine*. <https://leclab.wixsite.com/spatial/post/pansharpening-sentinel-2-imagery-in-google-earth-engine>
- Hernández, C., Escribano, J., & Tarquis, A. (2014). Comparación del Índice de Vegetación de Diferencia Normalizada obtenido a diferentes escalas en pastos de Dehesa. *Pastos y PAC 2014-2020*, 121–128.
- Hernández, F., & Martí, Y. (2007). *Conocimiento organizacional: la gestión de los recursos y el capital humano*. *14*, 26.

- Hernández, J., & Montaner, D. (2009). Patrones De Respuesta Espectral. In C. Mena (Ed.), *Tecnologías Geoespaciales: Experiencias Aplicadas al Estudio y Gestión del Territorio*. SAF. [http://www.gep.uchile.cl/Publicaciones/Hernández & Montaner 2009 SAF.pdf](http://www.gep.uchile.cl/Publicaciones/Hernández%20&%20Montaner%202009%20SAF.pdf)
- I.B., S., R., A., & J.A., A. (1999). Aplicación de sensores remotos en la detección y evaluación de plagas y enfermedades en la vegetación. *Teledetección. Avances y Aplicaciones. VIII Congreso Nacional de Teledetección*, 64–67. <http://www.aet.org.es/congresos/viii/alb16.pdf>
- Ingeoexpert. (2020). *Clasificaciones de imágenes de satélite*. <https://ingeoexpert.com/articulo/clasificaciones-de-imagenes-de-satelite/>
- INGEOMINAS. (2008). *Introducción a la Percepción Remota y sus Aplicaciones Geológicas*.
- Kaplan, G. (2018). Sentinel-2 Pan Sharpening—Comparative Analysis. *Proceedings*, 2(7), 345. <https://doi.org/10.3390/ecrs-2-05158>
- Koehrsen, W. (2017). *Understanding the Random Forest with an intuitive example*. <https://medium.com/@williamkoehrsen/random-forest-simple-explanation-377895a60d2d>
- León, Y. (2002). *Introducción a las Imágenes Satelitales*.
- Li, Y., Qu, J., Dong, W., & Zheng, Y. (2018). Hyperspectral pansharpening via improved PCA approach and optimal weighted fusion strategy. *Neurocomputing*, 315, 371–380. <https://doi.org/10.1016/j.neucom.2018.07.030>
- Liu, C., Frazier, P., & Kumar, L. (2007). Comparative assessment of the measures of thematic classification accuracy. *Remote Sensing of Environment*, 107(4), 606–616. <https://doi.org/10.1016/j.rse.2006.10.010>
- MADR. (2018). *Estrategia de Política Pública para la Gestión Integral de Riesgos Agropecuarios en Colombia*.
- MAE-MAGAP. (2015). *Protocolo metodológico para la elaboración del Mapa de cobertura y uso de la tierra del Ecuador Continental 2013-2014, escala 1:100.000*.
- Martínez Barbáchano, R., & Solís Miranda, G. A. (2018). Caracterización Espectral y Detección de Flecha Seca en Palma Africana en Puntarenas, Costa Rica. *Revista Geográfica de América Central*, 2(61), 349–377. <https://doi.org/10.15359/rgac.61-2.13>
- Millard, K., & Richardson, M. (2015). On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping. *Remote Sensing*, 7(7), 8489–8515. <https://doi.org/10.3390/rs70708489>
- Monterroso, M. (2013). Guía Práctica: Clasificación De Imágenes Satelitales. In *Software* (p. 31).
- Monterroso, M., Londoño, J., & Sansonov, S. (2018). Estimación del retroceso glaciario en los volcanes Nevado del Ruiz, Tolima y Santa Isabel, Colombia a través de imágenes ópticas y Din-SAR. *Dyna*, 85(206), 329–337. <https://doi.org/10.15446/dyna.v85n206.66570>
- Muñoz Aguayo, P. (2013). Apuntes de Teledetección: Índices de vegetación. In *Centro de información de recursos Naturales*. <https://doi.org/http://bibliotecadigital.ciren.cl/bitstream/handle/123456789/26389/Tema%20Indices%20de%20vegetaci%C3%B3n%20Pedro%20Mu%C3%B1oz%20A.pdf?sequence=1&isAllowed=y>
- Nonni, F., Malacarne, D., Pappalardo, S. E., Codato, D., Meggio, F., & De Marchi, M. (2018). Sentinel-2 Data Analysis and Comparison with UAV Multispectral Images for Precision Viticulture. *GI Forum*, 1, 105–116. https://doi.org/10.1553/giscience2018_01_s105
- Oñate de Mora, M. (2015). *Los Drones y sus aplicaciones a la ingeniería civil*.
- Oracle. (2020). *Inteligencia Artificial (IA) de Oracle: ¿Qué es el aprendizaje automático?* <https://www.oracle.com/co/artificial-intelligence/what-is-machine-learning.html>
- Ordoñez, A., & Serna, J. (2015). *Análisis Superficial y Multitemporal de Imágenes Landsat 7 Etm+*

- y Landsat 8 Oli Tirs En El Proyecto Carbonífero La Luna Entre Los Años 2001 Y 2015. Universidad de Manizales.
- Paruelo, J. M. (2008). La caracterización funcional de ecosistemas mediante sensores remotos. *Ecosistemas*, 17(3), 4–22. <http://www.revistaecosistemas.net/articulo.asp?Id=575>
- Pasqualotto, N., Delegido, J., Pezzola, A., Winschel, C., & Moreno, J. (2019). Estimación del contenido de clorofila a nivel de cubierta (CCC) en cultivos: Comparativa de índices de vegetación y el producto de nivel 2A de Sentinel-2. *XVIII Congreso de La Asociación Española de Teledetección*, September, 24–27.
- Paz Cerda, J. F. (2004). *Análisis multitemporal del recurso forestal en la Mancomunidad del Yegüare con el uso de imágenes satelitales LANDSAT 7*.
- Paz Pellat, F., Romero Sánchez, M. E., Palacios Vélez, E., Bolaños González, M., Valdez Lazalde, J. R., & Aldrete, A. (2014). Alcances Y Limitaciones De Los Índices Espectrales De La Vegetación: Marco Teórico. *Terra Latinoamericana*, 32(3), 177–194.
- Pérez Gutiérrez, C., & Muñoz Nieto, Á. L. (2006). *Teledeteccion: Nociones y Aplicaciones*. https://books.google.es/books?hl=es&lr=&id=SfrGxbO1DT0C&oi=fnd&pg=PA1&dq=historia+de+la+teledeteccion&ots=pEmi0B2gB-&sig=x_w5WFhCJmCJlJJDqcuXw3RjLE#v=onepage&q=historia+de+la+teledeteccion&f=false
- Redacción APD. (2019). *¿Cuáles son los tipos de algoritmos del machine learning?* <https://www.apd.es/algoritmos-del-machine-learning/>
- Reuter, F. (2009). Principios de teledetección. In *Serie Didáctica No 33. Catedra de Teldetección y Cartografía*.
- Robles, J., Docampo, M. de la luz, Sanz, J., Rodríguez, S., Larnaga, F., Sánchez, L., Hernández, S., Clérigo, Z., & Ortiz, L. (2016). Determinación de biomasa en parcelas de cultivos herbáceos mediante cámaras ópticas elevadas por medio de vehículos aéreos no tripulados (UAV). *Investigación Cualitativa En Ingeniería y Tecnología*, 4, 95–103.
- Russo, C., Ramón, H., Alonso, N., Cicerchia, B., Esnaola, L., & Tessore, J. P. (2016). *Tratamiento Masivo de Datos Utilizando Técnicas de Machine Learning*. 131–134.
- Sancha Navarro, E. F. (2010). El estudio de los índices de vegetación como base para conocer las relaciones entre la vegetación y el clima. In S. de P. de la U. de Sevilla (Ed.), *Tecnologías de la Información Geográfica: La información Greográfica al Servicio de los Ciudadanos* (pp. 1095–1108).
- Sarría, F. A. (2006). *Sistemas de Información Geográfica*.
- Stuckens, J., Coppin, P. R., & Bauer, M. E. (2000). Integrating contextual information with per-pixel classification for improved land cover classification. *Remote Sensing of Environment*, 71(3), 282–296. [https://doi.org/10.1016/S0034-4257\(99\)00083-8](https://doi.org/10.1016/S0034-4257(99)00083-8)
- TELEDET. (2020). *Satélites: resolución espectral*. <http://www.teledet.com.uy/tutorial-imagenes-satelitales/satelites-resolucion-espectral.htm>
- Traganos, D., Aggarwal, B., Poursanidis, D., Topouzelis, K., Chrysoulakis, N., & Reinartz, P. (2018). Towards Global-Scale Seagrass Mapping and Monitoring Using Sentinel-2 on Google Earth Engine: The Case Study of the Aegean and Ionian Seas. *Remote Sensing*, 10(8), 1227. <https://doi.org/10.3390/rs10081227>
- Tristan, P. M., Wainschenker, R. S., & Doorn, J. H. (2008). Normalización de Imágenes Satelitales en el Análisis Multi-Temporal. *X Workshop de Investigadores En Ciencias de La Computación*, 1, 315–319.
- UNCPBA. (2012). *Curso de Sensores Remotos*.

- Yébenes Gómez, Á., & Giner Sotos, Á. (2010). Clasificación Digital De Imágenes Por Satélite. In *Comprobar Fecha*.
- Zhao, L., Shi, Y., Liu, B., Hovis, C., Duan, Y., & Shi, Z. (2019). *Finer Classification of Crops by Fusing UAV Images and Sentinel-2A Data*.
- Zurqani, H. A., Post, C. J., Mikhailova, E. A., Schlautman, M. A., & Sharp, J. L. (2017). Geospatial analysis of land use change in the Savannah River Basin using Google Earth Engine. *International Journal of Applied Earth Observation and Geoinformation*, 69(September 2017), 175–185. <https://doi.org/10.1016/j.jag.2017.12.006>