

Potencial de sensores remotos en la detección de fosas clandestinas

Potential of remote sensing in detection of mass graves

Potencial dos sensores remotos na detecção de sepulturas clandestinas

Andrés Alejandro Mejía López ^{a*} | Luis Hernán Ochoa Gutiérrez ^b

a <https://orcid.org/0009-0000-7614-7429> Universidad Nacional de Colombia, Bogotá D.C., Colombia

b <https://orcid.org/0000-0002-3607-7339> Universidad Nacional de Colombia, Bogotá D.C., Colombia

- Fecha de recepción: 04/06/2024
 - Fecha concepto de evaluación: 16/09/2024
 - Fecha de aprobación: 22/09/2024
- <https://doi.org/10.22335/rlct.v16i3.1973>

Para citar este artículo/To reference this article/Para citar este artigo: Mejía, A. & Ochoa, L. (2024). Potencial de sensores remotos en la detección de fosas clandestinas. *Revista Logos Ciencia & Tecnología*, 16(3), 30-45. <https://doi.org/10.22335/rlct.v16i3.1973>

RESUMEN

En Colombia, la identificación de fosas comunes se ha convertido en una labor necesaria para el esclarecimiento de la verdad del conflicto violento. La actividad se realiza manualmente en la mayoría de los casos; para identificar los sitios de enterramiento se usa como insumo la información suministrada por los victimarios o los relatos de familiares de las víctimas. Sin embargo, estos métodos, además de imprecisos, son extremadamente tediosos y costosos.

Para incrementar las posibilidades de éxito de dicha labor, en otros países se han experimentado métodos de búsqueda basados en resistividad geofísica, magnetometría o radar de penetración del suelo. Estos métodos solo pueden usarse en un área delimitada, a menudo pequeña, además de presentar retos si están en lugares de difícil acceso como ciénagas, pantanos o en lugares donde persiste el conflicto, lo que podría implicar que estén minados. Por consiguiente, existe la necesidad de implementar métodos que no impliquen acceso directo al terreno y que puedan abarcar áreas extensas, agilizando la búsqueda y reduciendo el costo en tiempo y en dinero. En este artículo se explora el uso de datos obtenidos a partir de sensores remotos para identificar cómo la salud de las plantas es afectada por el contenido de materia orgánica presente en las fosas comunes, correlacionando estas con el análisis de los índices espectrales NDVI, GNDVI y GCI.

Para obtener imágenes sobre el área experimental, donde se simuló tumbas con materia orgánica de origen animal y restos óseos humanos, se utilizó el sensor multispectral Sequoia Parrot, transportado en un UAV DJI Phantom 3 Advanced, con el cual se adquirieron imágenes en las longitudes de ondas Red, Green y nIR, para medir los índices de vegetación y estimar la salud de las plantas sobre y alrededor de las fosas. Los resultados obtenidos permitieron inferir los lugares donde se encuentran inhumados los cuerpos en el sitio experimental después de casi una década de haber sido creado el modelo experimental.

Palabras clave: sensores remotos, fosas comunes, índices espectrales, aprendizaje de máquina, imágenes multispectrales, ortoimagen.

*Autor de correspondencia. Correo electrónico: aamejial@unal.edu.co



ABSTRACT

The identification of mass graves in Colombia has become an essential part of uncovering the truth of the violent conflict. This activity is mostly done manually, and in order to identify burial sites the killer's or the family's input is necessary. However, these methods, aside from being unreliable, are extremely complicated and expensive.

Research has been carried out, in other countries, to increase the success of this work, based on geophysical resistivity, magnetometry or ground penetration radar. These methods can only be used in a restricted range, aside from the added difficulties when they are utilised in hard-to-reach places such as swamps, marshes, or places where active armed conflict exists, which could mean the presence of landmines. Hence the need to use long range methods that avoid direct contact, expediting searches and reducing the cost in time and money. This article explores the use of data obtained from remote sensing to identify how the health of vegetation is affected by the content of organic matter present in burial sites, correlating these with the spectral indices NDVI, GNDVI, and GCI.

To obtain images of the experimental area where the burial simulacrum was carried out, with animals and human bones as organic matter, a Sequoia Parrot multispectral sensor, carried by a UAV DJI Phantom 3 Advanced, was used to gather the images in the Red, Green and nIR wave frequencies in order to measure the vegetation indices and gauge the plants' health over and around the burial sites. The results obtained allowed to deduce the places where the experimental graves were located after over a decade of being buried.

Keywords: Remote sensing; Mass graves; Spectral index; Machine learning; Multispectral imaging; Orthophoto.

RESUMO

Na Colômbia, a identificação de valas comuns tornou-se tarefa necessária para o esclarecimento da verdade sobre o conflito violento. Na maioria dos casos, essa atividade tem sido realizada manualmente, ou seja, para encontrar os locais de sepultamento, as informações fornecidas pelos perpetradores ou os relatos dos parentes das vítimas são usadas como entrada. Entretanto, esses métodos não são apenas imprecisos, mas também extremamente lentos e caros. Para aumentar as chances de sucesso, métodos de busca baseados em resistividade geofísica, magnetometria ou radar de penetração no solo foram experimentados em outros países. Esses métodos só podem ser usados em uma pequena área e apresentam desafios se forem encontrados em locais de difícil acesso, como pântanos, lodaçais ou em locais onde o conflito persiste e pode até mesmo ser minado. Portanto, há necessidade de implementar métodos que não envolvam acesso direto ao solo e que possam cobrir áreas maiores, a fim de acelerar a busca, reduzindo o custo em tempo e dinheiro. Este trabalho explora o uso de dados de sensoriamento remoto para

identificar como a saúde das plantas é afetada pelo conteúdo de matéria orgânica nas valas comuns, correlacionando estas como a análise dos índices espectrais NDVI, GNDVI e GCI. Neste trabalho, o sensor multiespectral Sequoia Parrot, transportado em um UAV DJI Phantom 3 Advanced, foi usado para obter imagens de uma área experimental na qual foram simuladas sepulturas com matéria orgânica de origem animal e restos de esqueletos humanos. A partir disso, foram adquiridas imagens nas longitudes de onda Red, Green e nIR, para medir os índices de vegetação e estimar a saúde das plantas sobre e ao redor das sepulturas. Os resultados obtidos permitiram inferir os locais onde os corpos estão enterrados no local do experimento quase uma década após a criação do experimento.

Palavras-chave: sensoriamento remoto, valas comuns, índices espectrais, aprendizado de máquina, imagens multiespectrais, ortoimagem.

Introducción

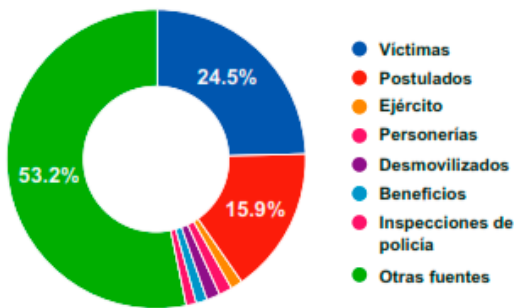
En Colombia se han registrado 8347566 víctimas a raíz del conflicto armado. Según los datos del "Registro Único de Víctimas" (RUV, 2017), se reporta un número indeterminado de desaparecidos.

Aunque en Colombia no existe acuerdo en la cifra, dos de las fuentes oficiales, el Sistema de Información Red de Desaparecidos y Cadáveres (SIRDEC, 2017) reporta 25212 desaparecidos (presuntamente forzados) a febrero de 2017 y el Registro Único de Víctimas (RUV, 2017) 46636 víctimas de desapariciones forzadas (directas) (véanse otras fuentes en la Figura 1).

Figura 1

Fuentes aportantes de información para la realización de exhumaciones

Fuentes de Información para la exhumación de cuerpos.
Grupo de exhumaciones FGN



Nota: Grupo de apoyo a exhumaciones Fiscalía General de la Nación (Equitas, 2015).

Lamentablemente, este fenómeno no es exclusivo de Colombia; más bien, parece un mal endémico de algunos países en vías de desarrollo. En Guatemala, por ejemplo, se estima en 200000 el número de desaparecidos producto del conflicto interno (FAFG, 2013; Kalacska et al., 2009), con patrones similares en otros países de América del Sur como Chile, Argentina, Perú, etc. (Kalacska et al., 2009). Los cuerpos de las víctimas en esos conflictos son frecuentemente enterrados en fosas comunes a menudo clandestinas, escondidas en la naturaleza en diferentes ubicaciones (Kalacska et al., 2009).

Esto ha propiciado que la sociedad civil busque la forma de identificar este tipo de estructuras, aunque en la mayoría de los casos ha recurrido a métodos tradicionales; usualmente, la prin-

cipal forma de abordar la búsqueda se hace de manera manual (de forma rudimentaria en muchos casos), guiados por la información de los victimarios (Blau et al., 2018; Silván-Cárdenas et al., 2021), o por los habitantes de las regiones en las cuales han sucedido las masacres (Molina, 2016).

A nivel internacional, se han aplicado métodos como la geomorfología forense, que tiene en cuenta las geoformas, su evolución, su relación con la geología y la biogeografía (Molina et al., 2015; Ruffell & McKinley, 2014); la termografía, que consiste en observar cambios en la temperatura del suelo para determinar anomalías (Davenport, 2001, 2018); métodos geofísicos que incluyen resistividad eléctrica, algunos apoyados en el uso del GPR (por sus siglas en inglés, *ground penetration radar*, radar de penetración del suelo), magnetometría y conductividad (Doro et al., 2022; Molina et al., 2016; Silván-Cárdenas et al., 2021), estos últimos limitados por las condiciones del terreno, accesibilidad del mismo, o la persistencia de actores armados en las zonas de búsqueda (Davenport, 2001; Doro et al., 2022; Molina, 2016).

En Colombia, el sistema judicial solo cuenta con un GPR, que empezó a usar de forma empírica y esporádica en el 2009, sin contar con ningún tipo de metodología o estandarización para su uso (Molina, 2016).

Recientemente, ha aumentado el uso de metodologías basadas en sistemas remotos, como escáneres láser "LiDAR" (Silván-Cárdenas et al., 2021), para localizar tumbas clandestinas. Dado que estas han demostrado un verdadero potencial, razón por la cual ha tenido un gran auge asociado a la actividad forense en países con recursos económicos y tecnológicos. En Colombia, se han desarrollado estudios experimentales y exitosos que incluyen el uso de sensores remotos (Molina et al., 2022) o imágenes multiespectrales (Mejía, 2024) para identificar fosas comunes. Es necesario proseguir con estos estudios para observar y entender su respuesta a lo largo del tiempo. El éxito de estos métodos es importante, si se tiene en cuenta que estos no son destructivos ni invasivos, lo que disminuye el riesgo de destruir la evidencia y protege al personal investigador en zonas de conflicto (Davenport, 2001; Silván-Cárdenas et al., 2021).

En la actualidad, las imágenes hiperespectrales (HSI, por sus siglas en inglés, *hyperspectral imaging*) han demostrado su utilidad en la localización de este tipo de estructuras. Estas se obtienen desde sensores aerotransportados o instalados a bordo de satélites, proporcionan más de 100 capas "bandas" con $\lambda < 10$ mm (Silván-Cárdenas et al., 2021); información que puede ser examinada y analizada selectivamente para detectar cambios sutiles en los espectros de reflectancia de la superficie (Kalacska & Bell, 2006). Asimismo, el uso de imágenes de sensores remotos para detectar los efectos de conflictos violentos ha experimentado un incremento dramático en los últimos años (Witmer, 2015), lo que, según Witmer, se ha dado en parte gracias a la mayor disponibilidad de imágenes satelitales.

A pesar de la utilidad y la eficacia de los sensores remotos aerotransportados, de los instalados en satélites comerciales, y que las HSI han mostrado una gran utilidad al localizar fosas comunes y entierros clandestinos, el costo de emplear estos sensores y tecnologías muchas veces resulta demasiado elevado (Evers & Masters, 2018) para algunas fuerzas del orden, ONG u organizaciones gubernamentales, como la Unidad de Búsqueda de Personas Desaparecidas en Colombia (UBPD).

En los últimos años se han hecho avances significativos en el desarrollo de vehículos aéreos no tripulados (UAV, por sus siglas en inglés, *unmanned aerial vehicle*), los cuales son cada vez más económicos (Ballarina et al., 2015; Urbanová et al., 2017), con mayor autonomía de vuelo y con más capacidad de carga (Evers & Masters, 2018); lo que a su vez ha permitido avances en la fotogrametría y hacer experimentos destinados a desarrollar técnicas más asequibles para identificar fosas comunes y sitios de enterramiento (Blau et al., 2018; Molina et al., 2022; Norton, 2019).

Otra motivación importante para hacer este trabajo, es el momento histórico que vive el país. Con los cambios que trae el posconflicto, es importante contribuir desde la academia, por lo que este artículo busca explorar el potencial de las imágenes multispectrales adquiridas a través de UAV para identificar fosas comunes antiguas (con más de dos años de haber sido creadas y de hasta diez años) como un paso

más para desarrollar una metodología asequible que permita agilizar la manera en que se realiza la búsqueda para encontrar los cuerpos.

El presente trabajo pretende abordar la detección de fosas comunes en Colombia, mediante el uso de imágenes tomadas con UAV, que contrastan la señal espectral de la vegetación y el suelo en el área de estudio.

Métodos

Para desarrollar el experimento se utilizaron las fosas simuladas desarrolladas en el trabajo llevado a cabo por Molina en el 2016, en el que se usaron cadáveres de cerdos domésticos para simular fosas clandestinas (véase Figura 2a), los cuales tienen una composición química, tamaño corporal, tejido, proporciones de grasa, tipo de piel y pelo similar al de los seres humanos (Molina, 2016).

En otra configuración, se utilizaron restos óseos humanos (véanse Figuras 2b y 2d), para representar tumbas clandestinas; esta configuración también se tuvo en cuenta al sobrevolar la zona, con la finalidad de hacer un estudio multitemporal y evaluar las respuestas de los diferentes tipos de sensores a lo largo del tiempo.

Figura 2

Fosas simuladas en el trabajo de Molina (2016)



Nota: (a) Fosa con cerdo; (b) Esqueleto humano; (c) Fosa de control; (d) Huesos incinerados y con la simulación de un decapitado.

El área experimental se ubica en las coordenadas 4°40'6.3" N, 74°12'33" W, a 2579 msnm, en el Centro Agropecuario Marengo, propiedad de la Universidad Nacional de Colombia, ubicado en el municipio de Mosquera (Cundinamarca) en el Km 14 de la vía que comunica a Bogotá con dicha población (véase Figura 3). El lugar de estudio es una zona rural con ecosistema considerado tropical con hierba y árboles, rodeado por construcciones del campus de la universidad. La geología del sector corresponde a terraza fluvio-lacustre de la formación Sabana,

con edad Pleistoceno medio a tardío (Molina, 2016); el clima presenta dos periodos secos y dos lluviosos con una precipitación aproximada de 1124 mm y una temperatura media que oscila entre los 12 °C-16 °C (IDEAM, 2017).

Se recolectaron muestras de vegetación y suelo para analizar la composición; luego, se utilizaron datos de pluviosidad y temperatura para incluir estas variables al estudio de resistividad eléctrica que se llevó a cabo (véase Figura 4).

Figura 3

Área de estudio

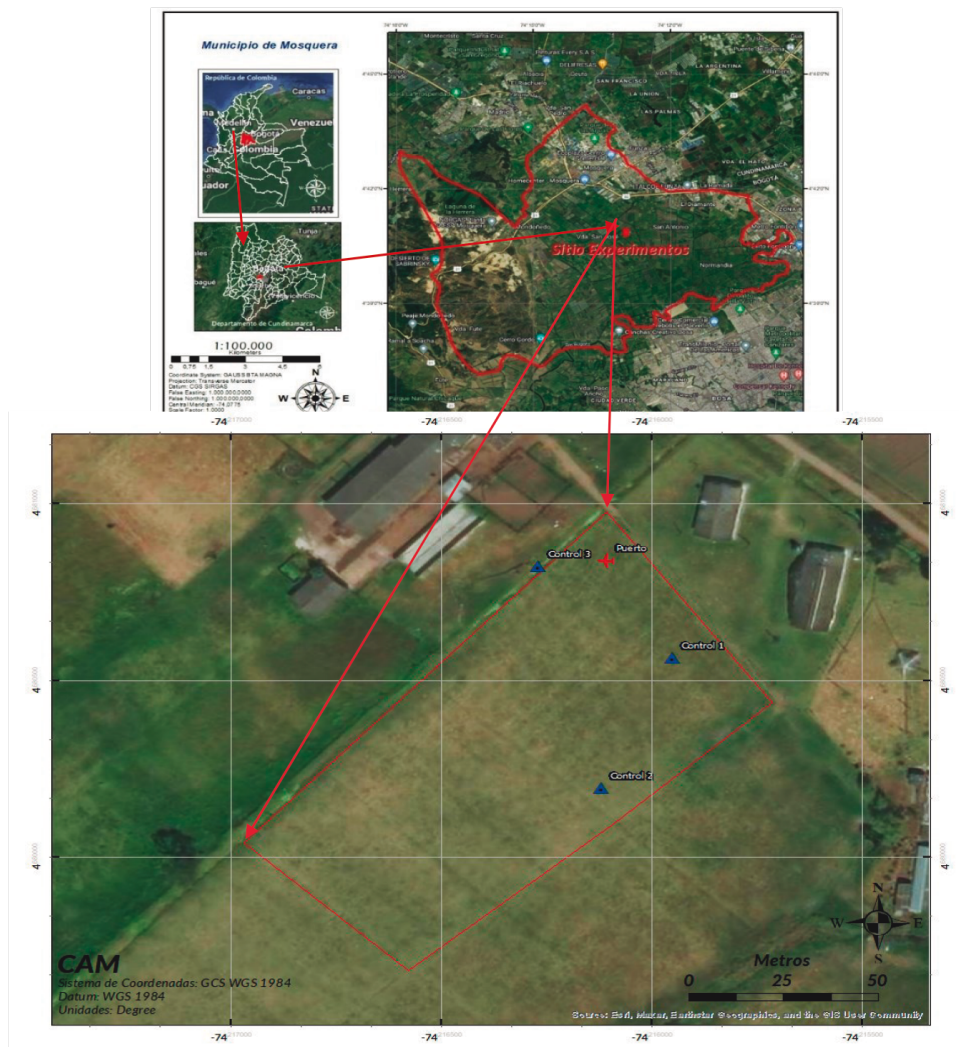
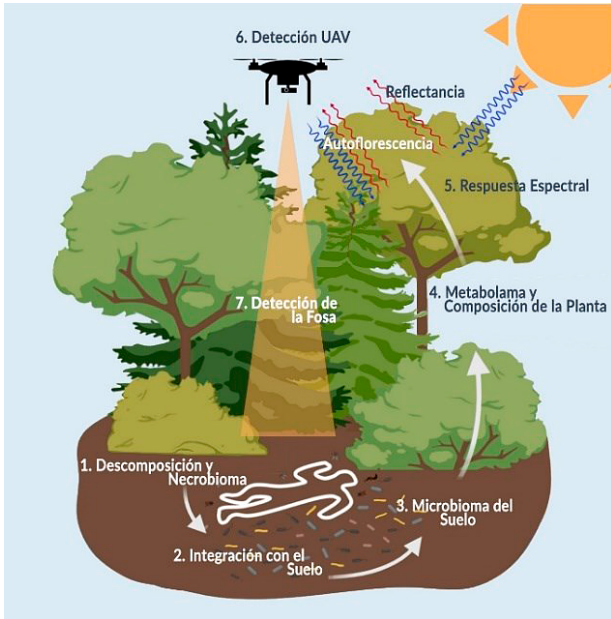


Figura 4

Modelo de detección remota de cuerpos en descomposición mediante UAV (Brabazon et al., 2020)



Nota: (1) Los cuerpos humanos en descomposición liberan compuestos. (2) Los compuestos se fusionan con el suelo y afectan su composición. (3) El microbioma del suelo se altera debido a los compuestos liberados por los cadáveres. (4) Las plantas que crecen sobre o cerca de los cuerpos pueden verse afectadas debido a estos cambios en el entorno. (5) Estos cambios pueden incidir en la salud de las plantas, lo cual puede ser medido a través de índices de vegetación que tienen en cuenta la respuesta espectral. (6) La recolección de datos de forma remota puede facilitar la recolección de datos e información. (7) Esta información lleva a la detección exitosa de la fosa común.

La adquisición se planeó en tres fases, con el fin de recrear el proceso actual para ubicar las inhumaciones que se llevan actualmente. En una primera fase, no se tuvo en cuenta la información exacta sobre dónde se encuentran las fosas (28 abril 2023); luego, se llevó a cabo una segunda misión en donde se sobrevoló el área con la información del investigador Martín Molina, quien nos cedió el modelo experimental (información de los perpetradores). Finalmente, en la tercera misión se adquirieron las imágenes para su procesamiento. El objetivo de este modelo es recrear las condiciones a las que se enfrenta un investigador al intentar detectar sitios de enterramiento.

Para adquirir las imágenes se usó un sensor multispectral Sequoia Parrot (véase Figura 5), el cual tiene una resolución por cada banda de 1.2 Mpx "definición: 1280 × 960 pixeles" y una

resolución radiométrica de 16 bits. El sensor adquiere cuatro bandas espectrales con anchos de banda para el rojo "Red: λ 640-680 nm", el verde "Green: λ 530-570 nm" y el infrarrojo cercano "nIR: λ 770-810 nm"; 10 nm para el borde del rojo "Red edge λ 730-740 nm". El sensor está equipado con GPS y sensor de luminosidad solar.

Figura 5

Sensor multispectral Sequoia Parrot



Figura 6

UAV DJI Phantom 3 Advanced



El sensor se instaló en un UAV DJI Phantom 3 Advanced, el cual sobrevoló a una altura de 30 m sobre el sitio de estudio (véase Figura 6). Durante el sobrevuelo se registraron puntos de control para corregir los datos GPS.

El vuelo para adquirir las imágenes (aerotriangulación) se realizó el 28 de abril de 2023; las siguientes misiones se planificaron para el

18 y 31 de mayo de 2023. Con el fin de hacer las correcciones, se toman imágenes del panel de referencia Spectralon instalado en el suelo para calibrar la radiancia y reflectancia de las imágenes. Para georreferenciar las imágenes, se usó una antena receptora del sistema global de navegación (GNSS, por sus siglas en inglés).

La adquisición de datos se planeó en tres fases, con el fin de recrear el proceso actual para ubicar las inhumaciones que se llevan actualmente:

1. En el primer sobrevuelo, no se tuvo en cuenta la información exacta sobre la ubicación de las fosas.
2. Sobrevuelo con la información aportada por el investigador Molina, que corresponde a la "información de los perpetradores".
3. Sobrevuelo de un área específica en el lote 5 de Marengo para el procesamiento.

Para la obtención de la información se tuvieron en cuenta las bandas *Red*, *Green* y *nIR*, pues son esenciales para calcular los índices espectrales, con el fin de medir o calcular la salud de la vegetación. Es importante tener en cuenta que la ubicación se basa en la identificación de los patrones de salud de la vegetación que crece alrededor de los sitios de inhumación.

La preparación del proceso fotogramétrico se llevó a cabo mediante los siguientes pasos:

1. Se prepara el vuelo usando "Litchi® for DJI Drones" insertando el área deseada que se va a explorar.
2. Se distribuyen y establecen cuatro puntos de control y se georreferencian para poder realizar la corrección geográfica de las imágenes.
3. Se toman imágenes del Spectralon para calibrar la radiancia y reflectancia de las imágenes.

El helipuerto del UAV DJI Phantom 2 se tomó como una referencia adicional a los puntos de control.

Se hizo una corrección radiométrica y geométrica para utilizar las imágenes adquiridas, con el fin de eliminar las variaciones de la inten-

sidad de la luz, la influencia de la atmósfera y los errores de calibración en el sensor. Para ello, se convierte la radiancia en reflectancia usando los valores de un panel de reflectancia conocida para determinar la relación entre la radiancia medida y la reflectancia conocida.

Para calcular los índices, se realizó primero la corrección radiométrica y geométrica de las imágenes adquiridas. Para ello, se usó el software PIX4Dfields® y Python3 "usando las siguientes librerías y software: *numpy*, *openCV*, *matplotlib*, *exiftool* + *pyexiftool*, *scikit-image*, *zbar* + *pyzbar*, *pysolar*, *pandas*, *mapboxgl*" para corregir los efectos electro-ópticos del sensor y la luz incidente del momento de captura.

Luego, se construyó un ortomosaico para cada banda, usando el software Agisoft® Metashape, teniendo en cuenta los puntos de control tomados *in situ* para anclar y corregir los datos tomados por el vuelo; cada banda se reproyectó para que tuvieran el mismo tamaño y la misma escala de pixel para poder calcular los índices de vegetación, esto se hizo con el software QGIS® en su versión 3.30.2.

El ortomosaico se realizó para cada una de las bandas de interés mediante el software Agisoft® Metashape 2.0.2. Para ello se siguieron los siguientes pasos:

1. Se importan las imágenes en el proyecto.
2. Se orientan o alinean las imágenes para calcular su posición y orientación en el espacio.
3. Se genera una nube de puntos 3D a partir de las imágenes alineadas.
4. En el modelado de superficie se genera una malla 3D a partir de la nube de puntos densa.
5. Se crea el ortomosaico utilizando la malla 3D y las imágenes originales (véanse Figuras 7 y 8).
6. Se guarda y exporta el ortomosaico en formato TIFF.
7. Teniendo en cuenta los puntos de control tomados *in situ*, se introduce georreferenciación de las bandas para anclar las fotos y corregir los datos del vuelo.

Figura 7

Imágenes del sensor RGB orientadas para construir el ortomosaico

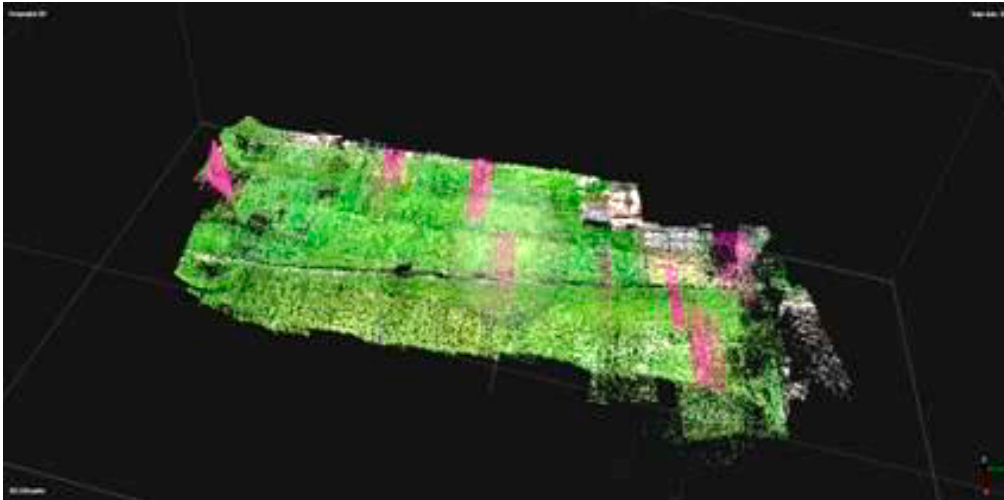


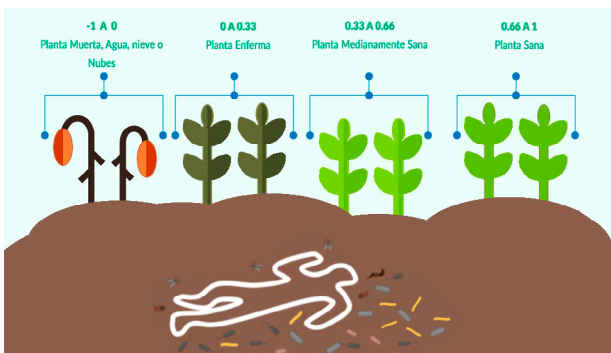
Figura 8

Malla creada con las imágenes RGB



Figura 9

Rangos NDVI. Tomados de Auravant (2023)



En este estudio, para calcular los índices espectrales se tomaron en cuenta los índices que permitieran identificar el estrés en las plantas que están creciendo sobre y alrededor de las fosas comunes:

$$NDVI = \frac{nIR - Red}{nIR + Red} \quad (1)$$

El índice diferencial de vegetación normalizada (NDVI, por sus siglas en inglés), desarrollado por Rouse et al. (1974), permite medir el contraste entre la máxima absorción de la clorofila (Chl) en la banda roja y la reflexión máxima en la banda infrarroja (Wu et al., 2010); se usa

para medir la salud y la densidad de la vegetación dentro del área de estudio. La vegetación saludable, rica en Chl, absorberá la luz roja y reflejará mayor cantidad en el espectro nIR (1). Los rangos de valores del NDVI van del -1 "usualmente atribuidos al agua" al 1 "verde intenso", pasando por el cero (0) "usualmente tomado como suelo desnudo" (véase Figura 9) (Molina et al., 2022).

$$\text{GNDVI} = \frac{nIR - \text{Green}}{nIR + \text{Green}} \quad (2)$$

Gitelson et al. (1996) definieron el índice diferencial de vegetación normalizada verde (GNDVI, por sus siglas en inglés) y asociaron las bandas del infrarrojo cercano y la banda verde del espectro electromagnético (495-570 nm). Este índice tiene una mayor sensibilidad al contenido de Chl en la vegetación, lo que permite establecer una estimación más precisa de la concentración del pigmento en las plantas. Este índice es comúnmente utilizado para determinar el consumo de agua y nitrógeno de la cubierta vegetal. El NDVI es medido en un rango de -1 a 1 en donde los valores cercanos a -1 y cero (0) indican agua y suelo, respectivamente.

$$\text{GCI} = \frac{nIR}{\text{Green} - 1} \quad (3)$$

El índice de clorofila (GCI, *green coverage index*) relaciona la banda infrarroja con la banda verde y permite, al igual que los anteriores, monitorear la salud de la planta, al estimar la cantidad de Chl; propuesto por Gitelson et al. (2003). El uso de la banda verde demostraba un incremento significativo para evaluar el contenido de Chl en las plantas al ser esta particularmente sensible a su variación.

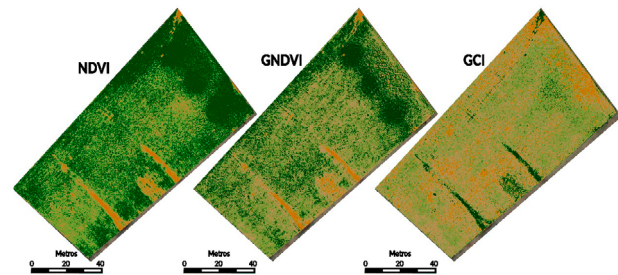
Estos índices se utilizan para medir los niveles diferenciales de Chl en las plantas expuestas al estrés. Estos niveles pueden determinarse en las características espectrales de las plantas, especialmente en las bandas relacionadas (Silván-Cárdenas et al., 2021). Por otra parte, si consideramos que un humano tiene alrededor de 2.6 kg de nitrógeno, "mucho del cual se convertirá en amonio durante la descomposición de los cadáveres" (Brabazon et al., 2020). Este

nitrógeno está estrechamente relacionado con el contenido de Chl en las hojas de las plantas. Si consideramos el tamaño de una fosa común y la cantidad de cuerpos en esta, la cantidad de nitrógeno liberado superará por mucho la recomendada para fertilizar (Brabazon et al., 2020), lo que afectará la señal espectral de la vegetación.

Los resultados que se muestran en la Figura 10, se calcularon usando las bandas *nIR*, *Red* y *Green* obtenidas con el sensor Sequoia Parrot. La clasificación se realizó en cuatro segmentos utilizando rupturas naturales "Jenks", para que los valores similares se agrupen mejor y maximizar la diferencia entre los valores. Al calcular y representar el índice NDVI, se pueden observar tres focos de interés importantes en los que se puede apreciar una señal más intensa. A pesar de que la señal es clara, se puede confundir un poco con la vegetación circundante. En el índice GNDVI el contraste es mejor que en el anterior; por último, con el GCI, aunque se puede identificar el contraste, no es claro.

Figura 10

Identificación fosas clandestinas



Para hacer la clasificación de los índices espectrales y el análisis de diferentes bandas de forma masiva, que permita potenciar dicha clasificación al depurar las mejores variables para crear un modelo, se aplica un proceso con *machine learning* (ML) a los datos recogidos en los siguientes pasos:

1. Preparación de datos

Para la construcción de la pila de imágenes y definición de archivos se crea una lista (*image_stack_resample*) que contiene variables de imágenes reproyectadas correspondientes a diferentes bandas espectrales e índices vegetativos (*Red*, *Green*, *nIR*, GCI, NDVI, GNDVI).

A su vez, se forma una lista (*file_list*) con las rutas de los archivos de estas imágenes, las cuales son almacenadas en un directorio especificado por la variable *fólder*.

La lectura y configuración de metadatos utiliza la librería *rasterio*. Se lee la primera imagen para extraer los metadatos que se emplearán en la creación de un archivo TIFF apilado y los metadatos se actualizan para reflejar la cantidad de bandas correspondientes al número de archivos. Finalmente, se da la escritura en la pila de imágenes, donde se escribe cada banda de cada archivo dentro del archivo TIFF de salida denominado *stack.tif*, respetando el orden y los metadatos definidos.

2. Procesamiento de datos

Seguidamente, se visualizan las dimensiones de cada banda y se reorganizan las matrices de las bandas para adecuarlas a los análisis futuros. Las imágenes en la pila son muestreadas utilizando puntos de un archivo de formas (*shapefile*), y los valores extraídos se almacenan en un *DataFrame* para su análisis y se normalizan los datos obtenidos del muestreo para prepararlos para el análisis de aprendizaje automático, utilizando técnicas como la estandarización o la normalización mín-máx, según se necesite.

3. Modelado predictivo

Para llevar a cabo el modelado hay dos etapas: la primera es el entrenamiento, usando varios modelos (SVM, árboles de decisión, *random forest*) con los datos preparados; luego se realiza una búsqueda en cuadrícula para optimizar los hiperparámetros y se evalúan los modelos basándose en su precisión utilizando datos de prueba. La segunda etapa es la de evaluación y visualización, en donde los modelos son evaluados y se presentan las métricas de desempeño. Se utilizan visualizaciones como la matriz de confusión para interpretar los resultados del modelo más efectivamente.

4. Aplicación de modelos a datos completos

Finalmente, se clasifican las imágenes, aplicando el modelo entrenado a la totalidad de la imagen para clasificar cada pixel con base en los patrones aprendidos, y los resultados se

reestructuran de nuevo en formato de imagen para su visualización final.

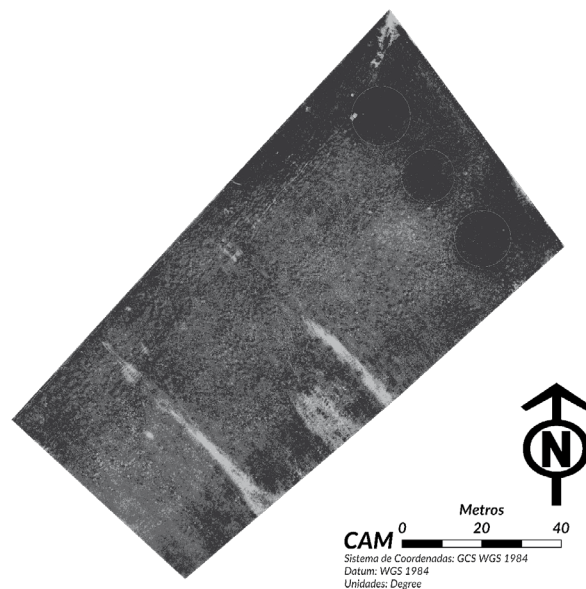
El código abarca una cadena de procesamiento de imágenes, desde la entrada y preparación de datos hasta el análisis predictivo mediante técnicas de aprendizaje automático, concluyendo con la aplicación del modelo a un conjunto de datos completos. Estos pasos reflejan un flujo de trabajo integral en el procesamiento y análisis de datos geoespaciales para aplicaciones de teledetección o similares, demostrando así la potencia de combinar técnicas de GIS con ML.

Resultados

Con la metodología planteada se pudieron crear cuatro ortomosaicos con cada uno de los índices propuestos, que permitieron realizar una comparación de los índices y variables, como se precisa en la Figura 11.

Figura 11

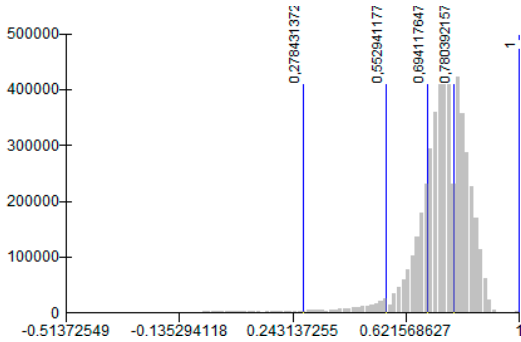
NDVI calculado del área de interés



NDVI calculado en el Centro Agropecuario Marengo en el sitio donde se encuentran las fosas simuladas. Se puede observar un área más oscura que tiene valores más cercanos a 1 con los colores más oscuros en la imagen, lo que corresponde a una vegetación sana. En el histograma (véase Figura 12) se puede observar la distribución de los valores.

Figura 12

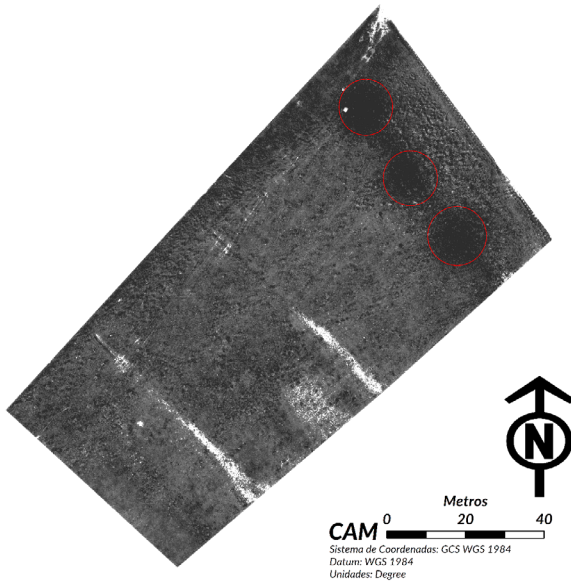
Histograma NDVI



Dado que los valores del NDVI van del -1 "usualmente atribuidos al agua" al 1 "verde intenso", pasando por el cero (0) "usualmente tomado como suelo desnudo", el histograma muestra un pico con una mayor concentración de datos alrededor del 0.78 (media de 0.74 y desviación estándar de 0.1), que indica áreas con vegetación saludable que absorben la luz roja y reflejan en el espectro del infrarrojo cercano (nIR).

Figura 13

GNDVI calculado del área de interés

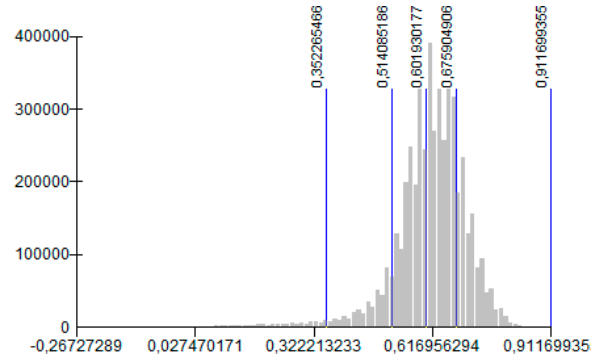


Con respecto al índice GNDVI calculado en el Centro Agropecuario Marengo sobre el sitio en el que se encuentran las fosas simuladas, al igual que en el NDVI, se puede observar un área más oscura que tiene valores más cercanos a 1, lo que corresponde a una vegetación sana o más vigorosa (véase Figura 13). En el

histograma (véase Figura 14) se muestra la distribución de los valores del GNDVI calculado.

Figura 14

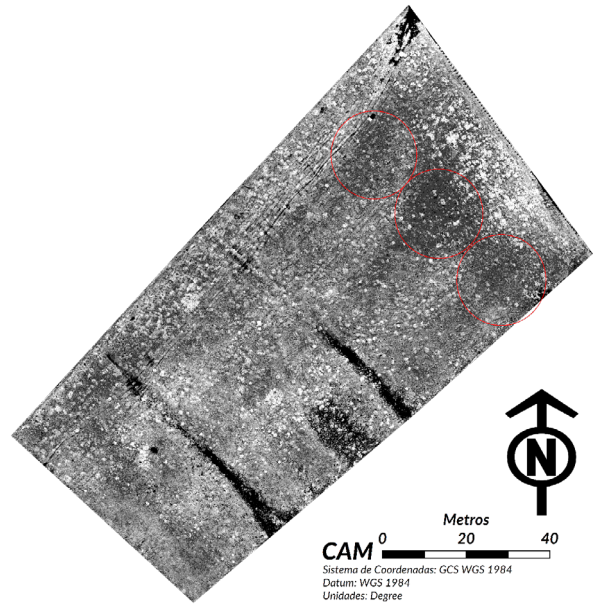
Histograma GNDVI



Los datos del histograma del GNDVI muestran un comportamiento similar al del NDVI con una distribución cercana a la normal, con un pico cercano a la media de 0.61 (desviación estándar de 0.08). Nuevamente, los datos más cercanos a 1 corresponden a una vegetación sana o vigorosa, dado que este índice tiene una mayor sensibilidad al contenido de clorofila en la vegetación.

Figura 15

GCI calculado del área de interés

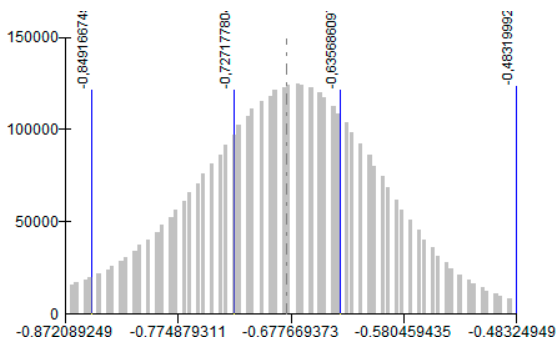


En el índice GCI se puede intuir el lugar donde la vegetación se muestra más saludable o vigorosa, al igual que en los otros índices esta se muestra con los colores más oscuros; en

este caso, los valores que se acercan a cero (0) muestran un mayor contenido de Chl (véase Figura 15).

Dado que los datos del histograma (véase Figura 16) parten del cálculo del GCI que resta 1 a la banda verde que divide a la banda infrarroja, los valores de GCI estarán centrados en torno al cero (0) e indican una mayor concentración de Chl (media -0.68 y desviación estándar de 0.1), que podría asociarse con una mayor sensibilidad a la variación en la banda verde.

Figura 16
 Histograma GCI



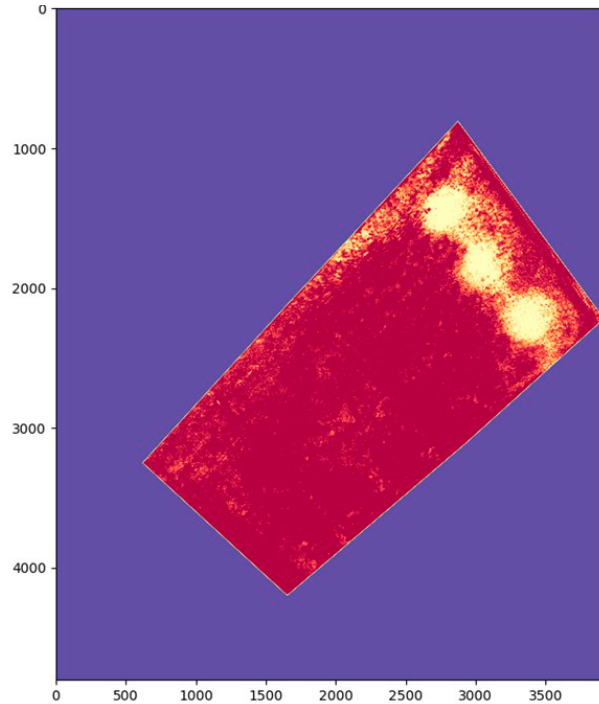
Del peso de las variables en el modelo de ML utilizado, se desprende que las variables que tuvieron mayor peso son la banda verde y el GNDVI (véase Tabla 1). Esto es coherente con la posición de algunos autores (Gitelson et al., 1996), quienes propusieron que la banda verde tenía mucha mayor sensibilidad para identificar la salud de las plantas.

Del ejercicio se desprende que el GNDVI es el índice más útil para identificar los sitios de inhumación, aunque no se puede descartar el uso del NDVI, ya que a simple vista también se pueden identificar los sitios de inhumación. Por el contrario, el GCI aunque involucre la banda verde, la dispersión de los datos no permite establecer a simple vista los lugares de inhumación.

En cuanto al uso de los métodos de ML, el resultado permitió iluminar los sectores donde se puede inferir la presencia de los cuerpos inhumados (véase Figura 17). En este se pudo ver la utilidad de los algoritmos basados en árboles de decisión, ya que estos se basan en la aleatoriedad y ante múltiples iteraciones pueden resaltar y escoger las variables de mayor

importancia para el modelo, a diferencia del SVM que no mostró un buen comportamiento para separar las variables y, por tanto, los resultados mostraron ser de poca utilidad.

Figura 17
 Clasificación con ML



Nota: En la imagen se pueden apreciar claramente y de forma diferenciada las anomalías en la vegetación. Se infiere que en los sitios colorados de amarillo se encuentran las fosas comunes experimentales. Se confirma que la descomposición de la materia orgánica crea hotspots alrededor de donde se realizó la inhumación.

Tabla 1
 Evaluación del peso de las variables

	Banda/Índice	Importancia
Banda	Green	30.67%
	Red	14.92%
	nIR	11.13%
Índice	GNDVI	19.67%
	NDVI	13.09%
	GCI	10.51%

En el modelo se evalúan y se comparan los diferentes índices y variables involucradas en la clasificación por RF; en esta se puede observar que la banda con mayor peso es la banda verde, lo que explicaría el mejor desempeño del GNDVI identificar las fosas comunes.

Discusión

En trabajos previos y citados en este documento, se ha concluido que la vegetación es un activo esencial para identificar sitios de inhumación, y este estudio concluye que la señal de la materia orgánica se puede observar aún después de casi diez años de haber sido creado el modelo experimental. Esto es importante en lugares en los que los conflictos se han extendido, como en Colombia, ya que las familias de las víctimas siguen esperando encontrar los restos de sus desaparecidos aun después de muchos años. Por tanto, es importante establecer si esa señal perdura en el tiempo mediante la toma de imágenes sobre un modelo de fosas experimentales que se construyó hace aproximadamente una década. Los resultados concluyen que el GNDVI es el índice más útil para identificar los sitios de inhumación, seguido del NDVI que permite a simple vista identificar las fosas y, por el contrario, el GCI que no permite establecer a simple vista los lugares de inhumación. Esto contrasta con lo concluido en la tesis de pregrado de González Gutiérrez, en el que establece que “el índice GCI se puede considerar como el más apto para identificar fosas comunes” (2023) a seis meses del entierro. Establecer el índice adecuado según el tiempo de la inhumación es necesario precisarlo para planear la búsqueda y obtener los mejores datos.

Además, los resultados han mostrado la utilidad de usar métodos de *machine learning* para asignarle peso a las diferentes variables. El algoritmo *Random Forrest* y CART mostraron que la banda verde, seguido del índice GNDVI, son las variables más importantes para identificar las fosas comunes (véase Tabla 1).

Al comparar los métodos con imágenes espectrales con las de geofísica de alta resolución, podemos observar que con las imágenes espectrales se pueden planear y realizar misiones sin importar el tipo de terreno, se pueden cubrir amplias áreas geográficas en poco tiempo y son más costo-efectivas al observar grandes extensiones de terreno. Asimismo, con las imágenes multiespectrales no importa el tipo o geoforma del suelo. Por otra parte, los métodos geofísicos son mejores con relación a la penetración del subsuelo, ya que tienen una

mayor resolución espacial que permite detectar características más pequeñas en el subsuelo en contraste con las imágenes multiespectrales que tienen un alcance solo sobre los elementos en la superficie.

Conclusiones

Luego de retomar el modelo experimental de Molina ubicado en el municipio de Marengo en la sede de la Universidad Nacional, se puede dar respuesta a la hipótesis planteada, con base en los resultados observados de cada uno de los índices y la clasificación hecha con ML apoyados en los algoritmos SMV, CART y RF. Este trabajo puede concluir que la señal de la materia orgánica se puede observar incluso después de casi diez años desde que se creó el modelo experimental, lo cual es importante en lugares en los que los conflictos se han extendido como en Colombia, ya que las familias de las víctimas siguen esperando encontrar los restos de sus desaparecidos, incluso después de muchos años.

Los índices de vegetación calculados demostraron ser útiles para identificar anomalías en la vegetación, lo cual es útil para establecer el lugar donde se encuentran ubicadas las fosas clandestinas.

Al comparar los tres índices, se observa una diferencia en su aplicación, notándose que el GNDVI ofrece un mayor grado de información para identificar las estructuras buscadas. Esto no solo es evidente a simple vista, sino que también se comprueba de manera empírica en la evaluación hecha en el modelo de ML empleado, donde se puede observar que al comparar los pesos de las variables, este es el índice con mayor calificación. Asimismo, la banda verde mostró un gran desempeño, ya que fue la banda con mejor calificación en el modelo.

Los modelos de ML empleados demostraron ser un método que facilita la identificación y la valoración de las variables, ya que al permitir combinar las bandas y los índices espectrales es mucho más claro el sitio de inhumación.

De los algoritmos utilizados, se obtuvo un mejor desempeño del CART y RF, logrando una precisión superior al 96% en los valores de

predicción. Esto se atribuye a los modelos basados en árboles de decisión que se apoyan en la aleatoriedad y, ante múltiples iteraciones, puede resaltar y escoger las variables de mayor importancia para el modelo.

El algoritmo SVM, por el contrario, no tuvo un buen comportamiento para separar las variables, por lo que con la mejor combinación de hiperparámetros solo alcanzó una precisión menor al 38%.

El uso de los índices espectrales demostró ser de utilidad para identificar la vigorosidad de la vegetación que puede estar cubriendo las fosas; ha demostrado ser especialmente útil el GNDVI al contrastar la banda Green contra el nIR; el índice GCI no permite hacer una inferencia clara para obtener información importante. Sin embargo, dado que aún faltan datos y estudios en la zona, los resultados son esperanzadores respecto al uso de los sensores remotos para ubicar la materia orgánica sepultada. Es necesario continuar el experimento para poder contrastar la información y obtener resultados concluyentes.

Finalmente, es importante anotar que el municipio de Mosquera, Cundinamarca, dónde se realizaron los experimentos, tiene unas especificaciones particulares de suelo y clima, siendo diferente a otras áreas del país donde probablemente hay inhumaciones, zonas con características más selváticas, con diferente pluviosidad o mayor humedad, lo cual se sugiere tener en cuenta para nuevos trabajos.

Referencias

- Auravant. (2023). Índices de vegetación y su interpretación: NDVI, GNDVI, MSAVI2, NDRE y NDWI. Auravant. Agricultura de Precisión. <https://www.auravant.com/blog/agricultura-de-precision/indices-de-vegetacion-y-como-interpretarlos/#quees-gndvi>
- Ballarina, M., Ballettia, C., & Guerra, F. (2015). *Action cameras and low-cost aerial vehicles in archaeology*. Proc. SPIE 9528, Videometrics, Range Imaging, and Applications XIII, 952813. <https://doi.org/10.1117/12.2184692>
- Blau, S., Sterenberg, J., Weeden, P., Urzedo, F., Wright, R., & Watson, C. (2018). Exploring non-invasive approaches to assist in the detection of clandestine human burials: Developing a way forward. *Forensic Sciences Research*, 3(4), 320-342. <https://doi.org/10.1080/20961790.2018.1493809>
- Brabazon, H., DeBruybe, J. M., Lenaghab, S. C., Li, F., Mundorff, A. Z., Steadman, D. W., & Stewart, C. N. (2020). Plants to remotely detect human decomposition? *Trends in Plant Science*, 25(10), 947-949. <https://doi.org/10.1016/j.tplants.2020.07.013>
- Davenport, G. (2018). *Remote sensing technology in forensic investigations: Geophysical techniques to locate clandestine graves and hidden evidence*. In T. Y. Group (Ed.), CRC Press.
- Davenport, G. C. (2001). Remote sensing applications in forensic investigations. *Historical Archaeology*, 35(1), 87-100. <https://doi.org/10.1007/bf03374530>
- Doro, K. O., Kolapkar, A. M., Bank, C.-G., Wescott, D. J., & Mickleburgh, H. L. (2022). Geophysical imaging of buried human remains in simulated mass and single graves: Experiment design and results from pre-burial to six months after burial. *Forensic Science International*, (335), 111289. <https://doi.org/10.1016/j.forsciint.2022.111289>
- Equitas. (2015). *Un radar para encontrarlos*. Cartilla sobre el Modelamiento Espacial y Estadístico Predictivo (MESP). <https://es.scribd.com/doc/289476071/Un-radar-para-encontrarlos>
- Evers, R., & Masters, P. (2018). The application of low-altitude near-infrared aerial photography for detecting clandestine burials using a UAV and low-cost unmodified digital camera. *Forensic Science International*, (289), 408-418. <https://doi.org/10.1016/j.forsciint.2018.06.020>
- FAFG (Fundación de Antropología Forense de Guatemala). (2013). *Asistiendo desde 1997 a las familias en la búsqueda e identificación de sus seres queridos desaparecidos*. FAFG. <https://www.fafg.org/>

- Gitelson, A. A., Gritz, Y., & Merzlyak, M. N. (2003). Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3), 271-282. <https://doi.org/10.1078/0176-1617-00887>
- Gitelson, A. A., Kaufman, Y. J., & Merzlyak, M. N. (1996). Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58(3), 289-298. [https://doi.org/10.1016/S0034-4257\(96\)00072-7](https://doi.org/10.1016/S0034-4257(96)00072-7)
- González Gutiérrez, L. G. (2023). *Identificación de fosas comunes mediante teledetección, aeronaves no tripuladas (UAV) e índices de vegetación* [Proyecto de Grado, Universidad de Cundinamarca]. Repositorio Digital Universidad de Cundinamarca
- IDEAM (Instituto de Hidrología, Meteorología y Estudios Ambientales). (2017). *Atlas climatológico de Colombia*. Imprenta Nacional de Colombia.
- Kalacska, M., & Bell, L. S. (2006). Remote sensing as tool for the detection of clandestine mass graves. *Canadian Society of Forensic Science Journal*, 39(1), 1-13. <https://doi.org/10.1080/00085030.2006.10757132>
- Kalacska, M., Bell, L., Sánchez-Azofeifa, A., & Celli, T. (2009). The application of remote sensing for detecting mass graves: An experimental case study from Costa Rica. *Journal of Forensic Sciences*, 54(1), 159-166. <https://doi.org/10.1111/j.1556-4029.2008.00938.x>
- Mejía, A. A. (2024). *Metodología para la identificación de fosas comunes a partir de imágenes multiespectrales* [Tesis de maestría, Universidad Nacional de Colombia]. Repositorio Institucional Universidad Nacional de Colombia
- Molina, C. M. (2016). *Metodología para la búsqueda de fosas a partir de la interpretación de anomalías en los datos obtenidos mediante la aplicación geofísica de alta resolución* [Tesis de doctorado, Universidad Nacional de Colombia]. Repositorio Institucional Universidad Nacional de Colombia
- Molina, C. M., Pringle, J. K., & Hernández, O. (2013). *Experiments to detect clandestine graves from interpreted high resolution geophysical anomalies*. <https://ui.adsabs.harvard.edu/abs/2013AGUSMNS21C..03M/abstract>
- Molina, C. M., Pringle, J. K., Saumett, M., & Evans, G. T. (2016). Geophysical monitoring of simulated graves with resistivity, magnetic susceptibility, conductivity and GPR in Colombia, South America. *Forensic Science International*, 261, 106-115. <https://doi.org/10.1016/j.forsciint.2016.02.009>
- Molina, C. M., Pringle, J. K., Saumett, M., & Hernández, O. (2015). Preliminary results of sequential monitoring of simulated clandestine graves in Colombia, South America, using ground penetrating radar and botany. *Forensic Science International*, 248, 61-70. <https://doi.org/doi=10.1016/j.forsciint.2014.12.011>
- Molina, C. M., Wisniewski, K., Heaton, V., Pringle, J. K., Ávila, E. F., Herrera, L. A., ..., & Baena, A. (2022). Monitoring of simulated clandestine graves of dismembered victims using UAVs, electrical tomography, and GPR over one year to aid investigations of human rights violations in Colombia, South America. *Journal of Forensic Sciences*, 67(3), 1060-1071. <https://doi.org/10.1111/1556-4029.14962>
- Norton, E. A. (2019). *A multi-temporal approach to using multispectral remote sensing for the prospection of clandestine mass graves in temperate environments*. [Tesis de doctorado, Bournemouth University]. Institutional Repository Bournemouth University

- Rouse, J., Haas, R., Schell, J., & Deering, D. (1974). Monitoring vegetation systems in the Great Plains with ERTS. (N. SP-351, Ed.) *Third Earth Resources Technology Satellite-1 Symposium, 1 Sect. A* (pp. 309-317). <https://ntrs.nasa.gov/api/citations/19740022614/downloads/19740022614.pdf>
- Ruffell, A., & McKinley, J. (2014). Forensic geomorphology. *Geomorphology*, 206, 14-22. <https://doi.org/10.1016/j.geomorph.2013.12.020>
- RUV (Registro Único de Víctimas). (2017). *Registro Único de Víctimas (RUV)*. Unidad para las Víctimas. <https://www.unidadvictimas.gov.co/registro-unico-de-victimas-ruv/>
- Silván-Cárdenas, J., Caccavari-Garza, A., Quinto-Sánchez, M., Madrigal-Gómez, J., Coronado-Juárez, E., & Quiroz-Suárez, D. (2021). Assessing optical remote sensing for grave detection. *Forensic Science International*, 329, 111064. <https://doi.org/10.1016/j.forsciint.2021.111064>
- SIRDEC (Sistema de Información Red de Desaparecidos y Cadáveres). (2017). *Registro Nacional de Desaparecidos*. Sistema de Información Red de Desaparecidos y Cadáveres. <https://siclico.medicinalegal.gov.co/consultasPublicas/>
- Urbanová, P., Jurda, M., Vojtíšek, T., & Krajsa, J. (2017). Using drone-mounted cameras for on-site body documentation: 3D mapping and active survey. *Forensic Science International*, 281, 52-62. <https://doi.org/10.1080/01431161.2015.1035412>
- Witmer, F. D. (2015). Remote sensing of violent conflict: Eyes from above. *International Journal of Remote Sensing*, 36(9), 2326-2352. <https://doi.org/10.1080/01431161.2015.1035412>
- Wu, C., Niu, Z., & Gao, S. (2010). Gross primary production estimation from MODIS data with vegetation index and photosynthetically active radiation in maize. *Journal of Geophysical Research*, 115(D12). <https://doi.org/10.1029/2009JD013023>