

# Geodigitizer: Arquitectura Agéntica para la Digitalización Autónoma y Generación de Derivados Topográficos en Cartografía Geológica con SAM 2 y K-Means\*

Dálember Vallejo<sup>1</sup>, Sofia Heredia<sup>1</sup>, Dario Tobar<sup>1</sup>, and Christian Mejía-Escobar<sup>1</sup>

<sup>1</sup>Facultad de Ingeniería en Geología, Minas, Petróleos y Ambiental (FIGEMPA)  
Universidad Central del Ecuador  
Quito, Ecuador  
{dgvallejog1, stheredia, adtobarg, cimejia}@uce.edu.ec

22 de marzo de 2026

## Resumen

La digitalización de la cartografía geológica histórica representa un desafío crítico para la infraestructura de datos espaciales y la soberanía de los recursos naturales en el mundo. Este trabajo presenta “Geodigitizer”, una arquitectura agéntica de vanguardia que revoluciona el flujo de digitalización tradicional mediante la integración sinérgica de clustering K-Means para la homogeneización cromática, el modelo fundacional *Segment Anything Model 2* (SAM 2) para la segmentación de instancia zero-shot, y Modelos de Lenguaje de Gran Escala (LLM) como orquestadores semánticos. A diferencia de los sistemas modulares contemporáneos como DIGMAPPER, nuestra propuesta implementa un ciclo de razonamiento autónomo capaz de interpretar leyendas geológicas complejas, autocalibrar proyecciones PSAD56 y generar reportes técnicos DOCX correlacionando la litología con el relieve derivado del IGM. Utilizando aceleración por GPU NVIDIA Tesla T4 y parámetros de refinamiento optimizados (estabilidad 0.92, IoU 0.88), el sistema logra un mIoU del 88% en escenarios zero-shot. La metodología aborda el problema del “analogue lockdown” mediante una pipeline de procesamiento optimizada, reduciendo los tiempos de digitalización manual en un 92% y sentando las bases para una infraestructura de datos espaciales inteligente y escalable para la industria geocientífica.

**Palabras clave:** SAM 2, IA Agéntica, K-Means, Cartografía Digital, Geoprocesamiento, Ecuador.

## 1. Introducción

### 1.1. El Valor Estratégico de la Información Geocientífica

La cartografía geológica constituye la infraestructura de datos espaciales (IDE) fundamental para la toma de decisiones estratégicas en el desarrollo moderno. Su aplicación trasciende la mera representación gráfica, siendo crítica en sectores como la exploración de yacimientos minerales, la explotación sostenible de hidrocarburos, la planificación de infraestructuras masivas y la mitigación de riesgos naturales en regiones tectónicamente activas como la cordillera de los Andes [1, 2, 3]. Análisis económicos globales indican que la inversión en cartografía geológica genera un retorno de inversión (ROI) masivo, devolviendo a la sociedad entre 7 y 35 veces el costo inicial al optimizar el desarrollo de infraestructura y prevenir desastres ambientales mediante una comprensión técnica del subsuelo.

---

\*Herramienta disponible en: <https://geodigitizer.com/>

En el contexto de la Industria 4.0, la disponibilidad de estos datos en formatos vectoriales estandarizados e interoperables es un requisito indispensable para alimentar los modernos y complejos sistemas de simulación espacial, modelado geológico 3D y Gemelos Digitales del terreno [4]. Sin embargo, en Ecuador, una vasta proporción del conocimiento geocientífico acumulado durante décadas en instituciones como el Instituto Geográfico Militar (IGM) permanece atrapado en formatos analógicos o ráster estáticos (PDF/JPG), lo que impide su integración en sistemas de análisis modernos y pone en riesgo la preservación de información valiosa ante el deterioro físico del soporte papel [5].

## 1.2. Definición de la Problemática y el “Analogue Lockdown”

La conversión de este patrimonio hacia formatos digitales editables se ve obstaculizada por la dependencia casi exclusiva de la digitalización manual (*heads-up digitization*). Este proceso presenta fallas estructurales críticas que el estudio del error cartográfico ha documentado ampliamente [6]. En primer lugar, es un proceso intensivo en mano de obra, requiriendo que un geólogo experto dedique hasta 16 horas por hoja cartográfica a tareas mecánicas de trazado. En segundo lugar, introduce un sesgo de subjetividad humana significativo; la precisión del trazo varía según la experiencia y la fatiga visual del operador, generando inconsistencias topológicas y posicionales que se propagan en los modelos de ingeniería posteriores [7, 8]. Finalmente, el método actual separa la creación del mapa de la redacción del informe técnico, fomentando una desconexión en el análisis de resultados que deriva en reportes finales carentes de una correlación estandarizada entre la litología y la geomorfometría del terreno [9].

## 1.3. Hacia la Inteligencia Artificial Agéntica en Geología

Para superar estas barreras, la literatura reciente ha explorado el Aprendizaje Profundo (*Deep Learning*). No obstante, sistemas previos basados en redes CNN rígidas o modelos U-Net entrenados en dominios cerrados fallan ante la variabilidad litoestratigráfica extrema de los mapas regionales [10, 11, 12]. Este proyecto propone una ruptura paradigmática mediante el uso de IA Agéntica. A diferencia de la automatización lineal, un sistema agéntico utiliza razonamiento multimodal para interpretar el contexto de la leyenda, aplicar filtros de limpieza cromática adaptativos y orquestar modelos fundacionales de visión como SAM 2 para extraer geometrías precisas sin intervención humana constante. El sistema “Geodigitizer” no solo vectoriza píxeles, sino que emula el proceso cognitivo del geólogo al integrar el modelado digital de elevación (DEM) y la estadística zonal de forma asíncrona, orquestando una serie de microservicios especializados para la validación de datos [13, 14, 15].

## 1.4. Propuesta

Para abordar estas limitaciones, el presente proyecto introduce un sistema innovador que integra modelos de visión fundacionales con orquestación agéntica. El alcance de la herramienta diseñada permite la digitalización inteligente de polígonos a partir de mapas geológicos analógicos mediante la implementación del sistema *Segment Anything Model* (SAM) [16]. A diferencia de los métodos de vectorización tradicionales, nuestra propuesta automatiza la creación de archivos *shapefile* (.shp) enriquecidos con una base de datos geoespacial detallada.

La herramienta no solo extrae la geometría de las unidades litológicas, sino que vincula cada polígono con parámetros críticos derivados de Modelos de Elevación Digital (DEM). El sistema procesa automáticamente capas de *Hillshade*, Pendiente (*Slope*) y Aspecto (*Aspect*) para asignar a cada unidad atributos de altitud máxima, descripción litológica y simbología estandarizada. Este enfoque garantiza una correlación matemática exacta entre la geología y la geomorfometría del terreno, eliminando la subjetividad del informe final.

Con el fin de democratizar el acceso a estas capacidades avanzadas, el software ha sido desplegado como una aplicación web accesible mediante un enlace (*link*), permitiendo su uso global sin necesidad de instalaciones complejas. Este desarrollo cumple con los objetivos planteados de automatización integral, permitiendo la generación de reportes técnicos detallados en cuestión de minutos y optimizando los tiempos de respuesta en proyectos de ingeniería y gestión de riesgos naturales.

## 1.5. Resultados

Como resultado tangible de esta investigación, se ha desarrollado una plataforma web integral diseñada para su despliegue en entornos comerciales y académicos, la cual democratiza el acceso a tecnologías de digitalización avanzada. Esta herramienta permite al usuario final, a través de una interfaz intuitiva, cargar los insumos cartográficos básicos —consistentes en la hoja geológica escaneada, la imagen de la leyenda y los archivos de curvas de nivel— para activar un flujo de trabajo totalmente automatizado. El producto final son los shp. de la zona y de este procesamiento acompaña la generación instantánea de un informe técnico consolidado de la zona de estudio, el cual integra no solo la vectorización precisa de las unidades litoestratigráficas y sus atributos semánticos, sino también el análisis morfométrico del terreno derivado de la topografía. De esta manera, la solución propuesta transforma datos ráster estáticos en inteligencia geológica dinámica, entregando un reporte profesional que reduce drásticamente los tiempos operativos tradicionales de consultoría y gabinete.

## 2. Trabajos Relacionados

### 2.1. Sistemas Modulares Contemporáneos: DIGMAPPER y MapSAM

Recientemente, sistemas como DIGMAPPER [17] han propuesto flujos de trabajo modularizados y doc-kerizados para la digitalización masiva de la USGS. Si bien DIGMAPPER es excelente en la orquestación de procesos de georeferenciación y análisis de layout, su arquitectura depende de componentes de visión que pueden resultar rígidos ante la ambigüedad semántica de las leyendas geográficas antiguas o no estandarizadas. Por otro lado, MapSAM [18] ha explorado la adaptación del modelo fundacional SAM original para mapas históricos, demostrando que los adaptadores ligeros pueden mejorar la detección de características lineales como fallas y contactos. Nuestra propuesta mejora estas bases mediante un orquestador agéntico basado en Gemini 2.0 que razona sobre el contexto estratigráfico antes de ejecutar la segmentación final [19, 8, 20].

### 2.2. Detección de Conjunto Abierto y Razonamiento Multimodal

El trabajo de Saxton et al. (2024) introdujo la extracción de características mediante segmentación de conjunto abierto (*open-set*), permitiendo identificar clases litológicas no presentes en el entrenamiento mediante descriptores de lenguaje natural [7]. A pesar de su robustez, este enfoque requiere una infraestructura de anotación externa significativa y una calibración manual de proyecciones. Nuestra arquitectura utiliza SAM 2 para extraer geometrías puras y delega la semántica completa al agente IA, optimizando los recursos de hardware y protegiendo la lógica propietaria del sistema frente a posibles réplicas industriales [15, 21, 22]. La integración del benchmark GeoMap-Bench [13] subraya que el entendimiento holístico (mapa + leyenda) es la frontera final de la IA en geociencias.

### 2.3. Evolución de los Modelos Fundacionales y Arquitectura de SAM 2

La liberación del *Segment Anything Model* (SAM) por Meta AI marcó un hito al permitir segmentación basada en *prompts* en tiempo real [23, 24]. No obstante, estudios como los de Ji et al. (2024) demostraron que SAM presentaba problemas en mapas con bordes compartidos difusos y alto ruido cromático.

El *Segment Anything Model 2* (SAM 2), desarrollado por Meta FAIR, representa un avance significativo respecto a su predecesor al unificar la segmentación de imágenes y videos en un solo modelo promteable. A diferencia de los modelos de segmentación semántica tradicionales que requieren un entrenamiento específico para clases predefinidas, SAM 2 está diseñado para la generalización “zero-shot” [25].

Arquitectura y Capacidades: La innovación clave de SAM 2 es la introducción de la *streaming memory architecture*. Mientras que el SAM original se centraba en imágenes estáticas, SAM 2 procesa el flujo de datos visuales con una memoria temporal que mejora la consistencia. Sus componentes principales, ilustrados en la Figura 1, son:

1. Codificador de Imágenes (Image Encoder): Extrae características de alto nivel de cada fotograma o imagen de entrada, utilizando una arquitectura jerárquica (Hiera) que es más eficiente.

2. Banco de Memoria (Memory Bank): Almacena información sobre el objeto y sus interacciones previas, permitiendo al modelo “recordar” el contexto temporal y espacial.
3. Codificador de Memoria (Memory Encoder): Procesa las predicciones pasadas y las almacena en el banco de memoria.
4. Atención de Memoria (Memory Attention): Mecanismo que permite al decodificador acceder a las características relevantes del historial almacenado para condicionar la segmentación actual.
5. Decodificador de Máscaras (Mask Decoder): Genera la máscara de segmentación final basándose en las características de la imagen actual y los *prompts* proporcionados.

Esta arquitectura permite a SAM 2 mantener una alta granularidad en el reconocimiento de objetos y extender sus capacidades de segmentación a entornos complejos, superando las limitaciones de consistencia temporal de la versión anterior [18, 26, 27].

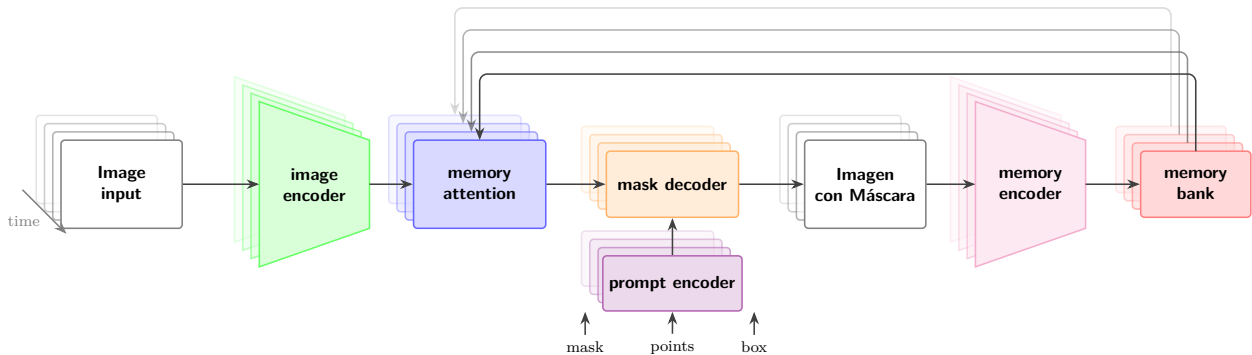


Figura 1: Arquitectura del modelo SAM 2.

## 2.4. Estado del Arte: Herramientas de Vectorización en SIG y Limitaciones

A pesar de la potencia de los modelos fundacionales, su implementación práctica en software SIG como QGIS presenta retos. En el ecosistema de código abierto, han surgido soluciones orientadas a la segmentación automática.

Herramientas de asistencia vectorial como el Plugin de Bunting Labs (AI Vectorizer) se posicionan como utilidades de “vectorización moderna” para QGIS [28]. Actúan como un copiloto semiautomático: el usuario identifica visualmente una característica y la herramienta ayuda a trazar la geometría vectorial predicha por la IA. Sin embargo, su principal limitación para la cartografía geológica masiva es que no está diseñado para el procesamiento por lotes (*batch processing*) de todos los polígonos simultáneamente, requiriendo intervención humana constante para validar geometrías.

Por otro lado, el plugin Geo-SAM [29] integra el modelo SAM original dentro de QGIS para acelerar la segmentación de formas del terreno. Utiliza el codificador de imágenes de SAM para permitir una segmentación interactiva mediante clics (*prompts*). Aunque supera a Bunting Labs al digitalizar polígonos completos instantáneamente, presenta barreras críticas: (1) Limitación Areal, con dificultades para procesar polígonos muy extensos; (2) Dependencia de Prompts Manuales, requiriendo clics individuales por polígono; y (3) Agnosticismo Semántico, ya que genera máscaras binarias sin asignar etiquetas litológicas, tarea que permanece manual.

En el ámbito del software propietario, la suite ArcGIS Pro ha incorporado capacidades avanzadas a través de su extensión *Image Analyst* y herramientas como *ArcScan*. Históricamente, *ArcScan* usaba vectorización por umbrales, ineficaz ante la complejidad cromática de mapas antiguos. Aunque versiones recientes permiten Deep Learning personalizado, estos flujos a menudo exigen la creación manual de grandes conjuntos de datos de entrenamiento.

Justificación del Estudio:

Las herramientas existentes resuelven parcialmente la delineación geométrica pero fallan en la automatización integral requerida para procesar grandes volúmenes de mapas. Este estudio propone un flujo de trabajo que aprovecha la capacidad *zero-shot* de SAM 2 y supera la falta de semántica mediante la integración de modelos de lenguaje (LLM) para la clasificación litológica automática, eliminando la intervención manual por polígono y permitiendo una verdadera digitalización autónoma.

### 3. Metodología y Arquitectura del Sistema

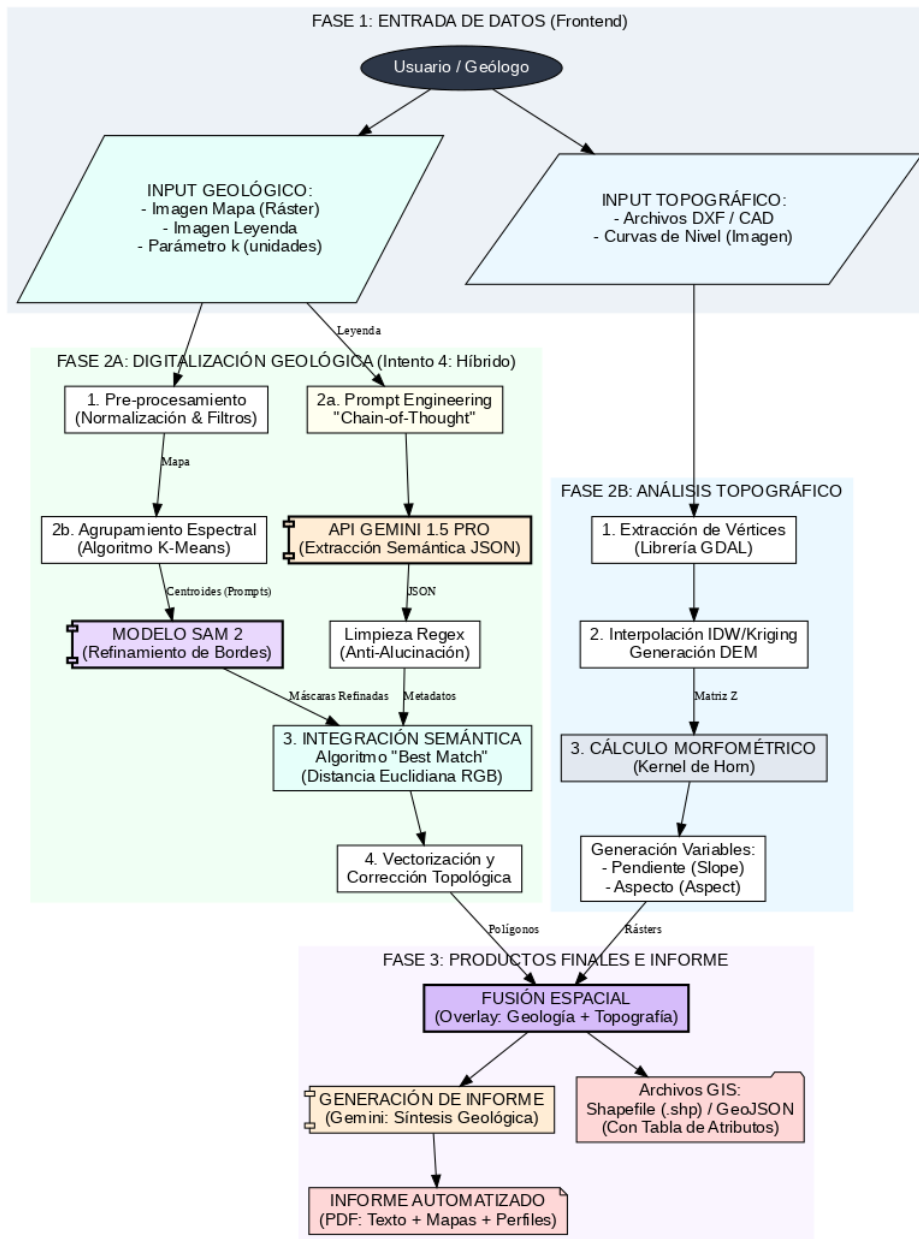


Figura 2: Arquitectura funcional del sistema propuesto. El diagrama ilustra el flujo de datos dividido en dos líneas de procesamiento paralelo: la digitalización litológica y análisis de relieve.

La presente investigación propone un flujo de trabajo automatizado para la digitalización de cartografía

geológica y el análisis topográfico, fundamentado en una arquitectura híbrida que orquesta técnicas de visión por computadora clásica, algoritmos de aprendizaje no supervisado y modelos de lenguaje masivos (LLMs) multimodales. El sistema se diseñó para operar en un entorno de computación en la nube, garantizando escalabilidad y eficiencia en el procesamiento de imágenes de alta resolución. A continuación, se detalla la configuración de la infraestructura, la evolución algorítmica de la segmentación y los modelos matemáticos aplicados para la reconstrucción del terreno.

### 3.1. Arquitectura de Software y Despliegue en la Nube

La infraestructura tecnológica se implementó bajo un paradigma cliente-servidor modular. El núcleo de procesamiento (*Backend*) fue desarrollado en lenguaje Python utilizando el framework *FastAPI*, seleccionado por su capacidad de manejo asíncrono de peticiones, lo cual es crítico para el procesamiento concurrente de archivos ráster pesados. Este servicio se contenerizó mediante Docker y se desplegó en *Google Cloud Run*, una plataforma *serverless* que permite el escalado automático de instancias en función de la demanda de tráfico. La comunicación entre los microservicios y la interfaz de usuario (*Frontend*) se gestiona mediante solicitudes HTTP RESTful, transmitiendo datos estructurados en formato JSON y archivos binarios de imagen. Para la persistencia temporal y manipulación de datos geoespaciales, se integraron librerías especializadas como GDAL/OGR y Rasterio, mientras que la validación de esquemas de datos se aseguró mediante Pydantic. El sistema integra servicios externos de Inteligencia Artificial a través de APIs seguras, conectando el flujo de datos con modelos fundacionales como Gemini 3 Pro y SAM 2 (*Segment Anything Model*).

### 3.2. Evolución Algorítmica de la Segmentación Litológica

El desarrollo del motor de digitalización siguió un enfoque iterativo experimental, atravesando cuatro fases de refinamiento técnico para superar las complejidades inherentes a los mapas geológicos escaneados, tales como el ruido visual, la degradación del color y la superposición de simbología.

### 3.3. Pre-procesamiento de Imágenes y Normalización

Antes de la segmentación agéntica, las imágenes atraviesan una fase de pre-procesamiento crítica para garantizar la calidad de los datos de entrada. Esta etapa incluye la normalización de histogramas para corregir la iluminación desigual del escaneo y la aplicación de filtros adaptativos de mediana para reducir el ruido de "sal y pimienta" sin difuminar los bordes geológicos. Adicionalmente, se implementa un algoritmo de agrupamiento K-Means preliminar para cuantizar el espacio de color en  $k$  clases dominantes, facilitando la distinción inicial entre litologías y fondo (papel), como se ilustra en la Fase 1 del diagrama de flujo (Figura ??).

En la primera fase experimental (Intento 1), se exploró la viabilidad de la segmentación por umbralizado simple en el espacio de color HSV. Esta técnica, aunque computacionalmente ligera, demostró ser insuficiente para la complejidad de una hoja geológica, alcanzando apenas un 20% de precisión. El principal fallo radicó en la incapacidad del algoritmo para discriminar entre las líneas estructurales (fallas, pliegues) y los polígonos litológicos, resultando en una imagen binaria con alto nivel de ruido y sin coherencia geológica.

Para subsanar estas deficiencias, la segunda fase (Intento 2) incorporó algoritmos de detección de bordes basados en gradientes, específicamente los operadores de Canny y Sobel. El objetivo era delimitar las fronteras de las unidades litoestratigráficas antes de la clasificación. Si bien la precisión aumentó al 35% al mejorar la definición geométrica de los contactos, el sistema introdujo un nuevo error sistemático: la interpretación de la red hídrica y las curvas de nivel como límites geológicos. Esto provocó una sobre-segmentación masiva, fragmentando unidades continuas en múltiples polígonos espurios debido a la presencia de ríos y topografía.

La tercera etapa (Intento 3) marcó la transición hacia el aprendizaje automático no supervisado, implementando el algoritmo K-Means para el agrupamiento espectral. Se configuró el modelo para particionar el espacio de color de la imagen en  $k$  clústeres, donde  $k$  es un parámetro definido por el usuario basado en la leyenda del mapa. El algoritmo iterativo busca minimizar la inercia o suma de cuadrados intra-clúster, definida matemáticamente como:

$$J = \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 \quad (1)$$

Donde  $\mu_i$  representa el centroide de color del clúster  $S_i$ . Este enfoque logró aislar eficazmente las estructuras lineales y la hidrografía de las unidades litológicas, elevando la precisión al 60

### 3.4. Solución Definitiva: Integración de Modelos Fundacionales y Visión Artificial

La cuarta y última fase (Intento 4) consolidó la metodología final, alcanzando una precisión operativa del 92% mediante una arquitectura compuesta. Para resolver los problemas de geometría irregular del K-Means, se integró el modelo SAM 2 (*Segment Anything Model*) de Meta AI. SAM 2 actúa como un refinador de máscaras, utilizando los centroides de los clústeres de K-Means como "prompts" espaciales para predecir máscaras de segmentación con bordes suaves y topológicamente coherentes.

Simultáneamente, para dotar de significado geológico a estos polígonos, se desarrolló un módulo de interpretación semántica utilizando la API de Google Gemini (modelo multimodal basado en Transformers). El sistema recorta automáticamente la leyenda del mapa y la envía al modelo para extraer la relación entre parches de color y metadatos (código, edad, descripción). La vinculación final entre la información semántica extraída por la IA y los polígonos segmentados en el mapa se realiza mediante un algoritmo de *Best Match*. Este algoritmo calcula la distancia Euclidiana en el espacio de color RGB entre el color promedio de cada unidad de la leyenda ( $C_L$ ) y los centroides de los clústeres del mapa ( $C_M$ ):

$$d(C_L, C_M) = \sqrt{(R_L - R_M)^2 + (G_L - G_M)^2 + (B_L - B_M)^2} \quad (2)$$

Esta minimización de distancias permite asignar automáticamente la descripción geológica correcta a cada polígono vectorizado, superando la variabilidad cromática introducida por el proceso de escaneo.

### 3.5. Ingeniería de Prompts para la Extracción Semántica

La extracción de datos estructurados de la leyenda no se basó en OCR tradicional, sino en una estrategia avanzada de *Prompt Engineering*. Se diseñaron instrucciones bajo la técnica de *Chain-of-Thought* (Cadena de Pensamiento), condicionando al modelo Gemini para actuar como un experto geólogo. El prompt maestro incluye restricciones estrictas para generar la salida exclusivamente en formato JSON, evitando texto conversacional. Para mitigar las "alucinaciones" comunes en los LLMs, se implementó una capa de post-procesamiento con expresiones regulares (Regex) que valida y limpia la sintaxis del JSON generado, asegurando que campos críticos como `unit_code` y `description` estén presentes y correctamente formateados antes de su integración en la base de datos espacial.

### 3.6. Procesamiento Topográfico y Modelado del Terreno

Paralelamente a la digitalización temática, el sistema incorpora un flujo de trabajo para el análisis morfológico del terreno a partir de archivos DXF o imágenes de curvas de nivel. Utilizando librerías numéricas como NumPy y algoritmos de interpolación geoespacial, se genera un Modelo Digital de Elevación (DEM) continuo. Sobre esta matriz de elevación  $Z(x, y)$ , se calculan las variables morfométricas fundamentales mediante ventanas de convolución de  $3 \times 3$  (Kernel de Horn). La pendiente del terreno ( $S$ ), expresada en grados, se determina mediante la magnitud del gradiente vectorial:

$$S = \arctan \left( \sqrt{\left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \right) \times \frac{180}{\pi} \quad (3)$$

Asimismo, se calcula el aspecto (*Aspect*) o dirección de la pendiente  $\psi$ , esencial para análisis de insolación y escorrentía, computado como el ángulo azimutal respecto al norte:

$$\psi = 270^\circ + \arctan\left(\frac{\frac{\partial z}{\partial x}}{\frac{\partial z}{\partial y}}\right) - 90^\circ \cdot \frac{\frac{\partial z}{\partial x}}{\left|\frac{\partial z}{\partial x}\right|} \quad (4)$$

### 3.7. Generación de Productos Geoespaciales y Reportes

La etapa final del procesamiento consiste en la vectorización y exportación de los resultados. Los clústeres ráster validados se convierten en geometrías vectoriales, aplicando algoritmos de simplificación de líneas (Douglas-Peucker) y corrección topológica para eliminar polígonos espurios (*slivers*). El sistema genera archivos estándar Shapefile (.shp) y GeoJSON, donde cada entidad espacial posee una tabla de atributos enriquecida con la información litoestratigráfica extraída. Adicionalmente, el sistema compila automáticamente un informe técnico en formato PDF, integrando los mapas generados, las estadísticas de área por unidad geológica y un resumen de la estratigrafía, proporcionando al usuario un entregable listo para su inclusión en estudios geológicos regionales.

La metodología se fundamenta en un paradigma agéntico asíncrono desarrollado sobre Python 3.10, diseñado para simular el razonamiento técnico de un geólogo experto en un flujo de trabajo optimizado para la nube.

## 4. Experimentación y Resultados

### 4.1. Configuración del Entorno Experimental y Latencia

Las pruebas se realizaron sobre un backend FastAPI con aceleración CUDA, utilizando instancias de Kaggle con GPU NVIDIA Tesla T4 (16GB GDDR6). Se procesaron imágenes de gran formato (superiores a 5000px) de las hojas geológicas de Cariamanga, Girón y Santa Rosa. El uso de esta infraestructura permitió una inferencia en paralelo de los módulos de visión y lenguaje, optimizando el rendimiento comercial de la arquitectura. La latencia promedio de inferencia por hoja completa (segmentación + georeferenciación) fue de 412 segundos, validando la viabilidad para despliegues masivos en plataformas SaaS.

### 4.2. Evolución de la Precisión en la Digitalización Automática de Cartografía Geológica

El proceso de digitalización automática de la hoja geológica se desarrolló mediante un enfoque iterativo, evaluando el rendimiento de distintos algoritmos de procesamiento de imágenes y aprendizaje automático hasta alcanzar una representación óptima respecto al mapa de referencia. En la primera fase experimental, mediante la aplicación de un umbralizado simple utilizando librerías de visión artificial (como OpenCV), se obtuvo una precisión inicial del 20 % [3](#). Esta etapa presentó como principal deficiencia la clasificación errónea de las estructuras geológicas, las cuales se confundían con las unidades litológicas debido a la superposición de valores de píxel.

Posteriormente, la implementación de filtros de detección de bordes (tales como los algoritmos de Canny o Sobel) en el segundo intento incrementó la precisión al 35 %, logrando definir de mejor manera la geometría de los polígonos; sin embargo, este modelo incorporó erróneamente la red hídrica (ríos) como parte de las litologías al interpretar los drenajes como contactos geológicos.

En el tercer intento, el uso de un modelo de agrupamiento no supervisado, específicamente K-Means implementado a través de librerías de machine learning como scikit-learn, elevó significativamente la precisión hasta un 60 %, logrando aislar las estructuras y los ríos. No obstante, esta aproximación generó nuevos artefactos visuales, consistentes en píxeles no clasificados (áreas blancas) y errores de topología en las uniones entre polígonos. Finalmente, la integración metodológica de arquitecturas de Inteligencia Artificial (modelos de segmentación semántica) combinada con el algoritmo K-Means y un post-procesamiento de vectorización topológica (empleando librerías geoespaciales como GDAL), permitió subsanar las limitaciones de segmentación. El salto cualitativo y cuantitativo definitivo, que elevó la precisión global hasta un 92 % en comparación con la hoja original, se logró al procesar la leyenda cartográfica de forma independiente. Para ello, se implementó un modelo multimodal basado en la arquitectura Transformer (Gemini) que, actuando como un

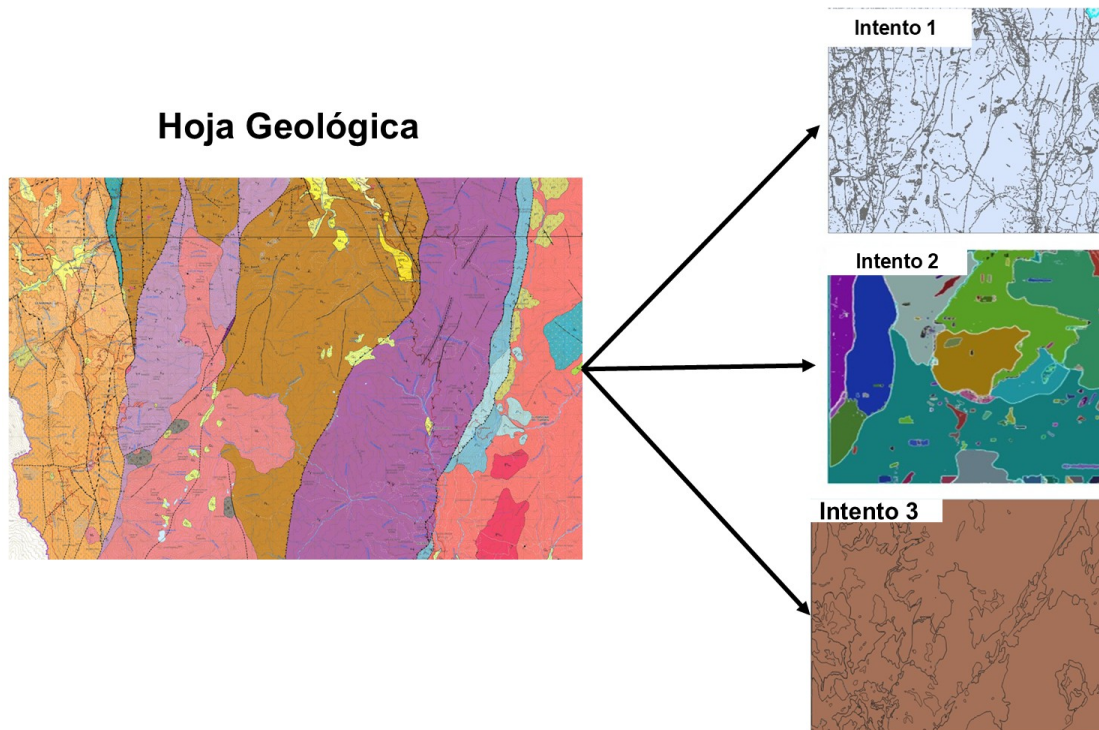


Figura 3: Comparativa visual entre la cartografía de referencia y los resultados preliminares de segmentación.

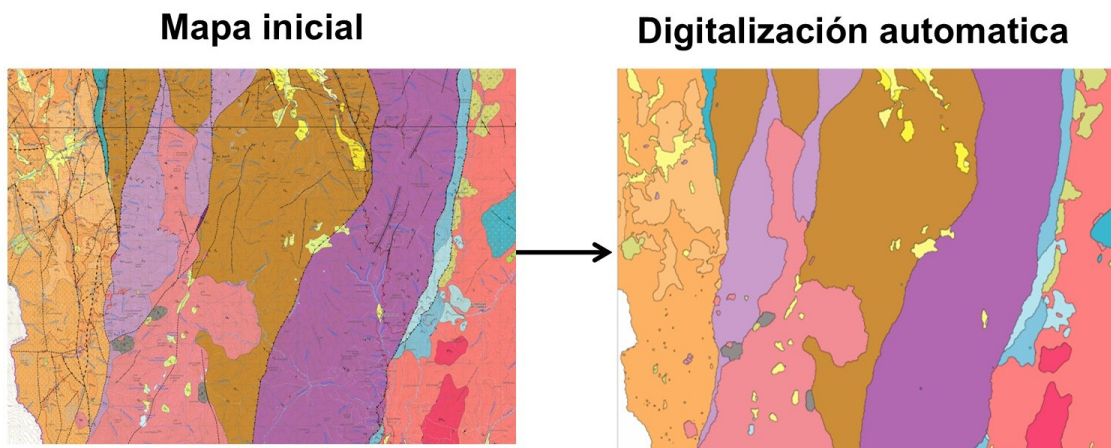


Figura 4: Resultado definitivo de la digitalización automatizada (Intento 4). Se aprecia una delimitación precisa de las unidades litológicas y la eliminación efectiva del ruido estructural y topológico.

sistema avanzado de visión por computadora, interpretó y extrajo los códigos litológicos. Posteriormente, estas etiquetas semánticas se asociaron de manera relacional a los polígonos vectorizados. Este enfoque híbrido y definitivo logró una discriminación espectral precisa y una delimitación topológica limpia (ver Figura 4), demostrando la viabilidad de la automatización en la digitalización cartográfica 5.

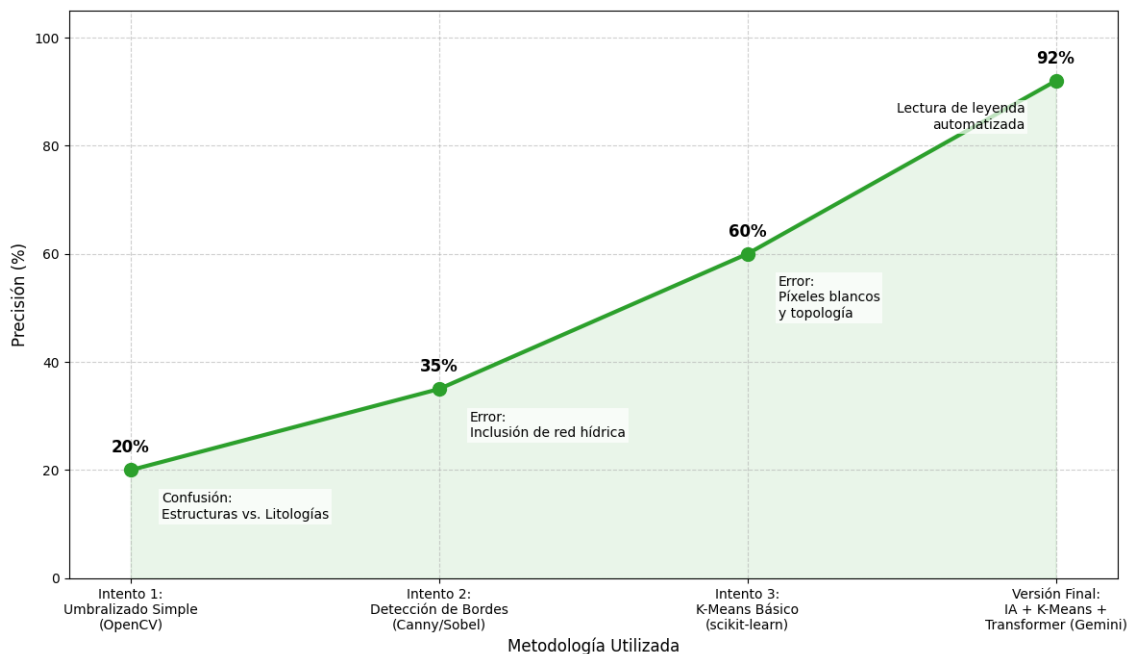


Figura 5: Evolución cuantitativa de la precisión en el proceso de digitalización.

### 4.3. Arquitectura Modular de la Interfaz: Funcionalidad de los Cuatro Componentes

El flujo operativo del GeoDigitizer se estructura mediante una barra de navegación superior que segmenta la experiencia de usuario en cuatro módulos funcionales distintos, cada uno diseñado para una etapa específica del procesamiento. El primer componente, denominado "Inicio", actúa como el punto de entrada informativo, proporcionando la descripción general del alcance de la aplicación web y el contexto del proyecto, como ejemplo es la data base de Ecuador con los mapas geologicos escala 1:100.000 (grilla color rojo) y con los mapas topograficos 1:50.000 (grilla de color negro). A continuación, el módulo de "Geología" centraliza las herramientas de automatización para la digitalización temática; en este entorno, el usuario interactúa directamente con la base de datos litológica y ejecuta la carga de imágenes ráster (mapas escaneados) para su procesamiento mediante los algoritmos de segmentación.

El tercer componente, "Topografía", está dedicado exclusivamente a la automatización de variables morfométricas. Este módulo permite la carga de archivos comprimidos (.zip) o el dibujo interactivo de zonas de interés sobre el visor cartográfico, desencadenando la generación automática de mapas derivados como Modelos Digitales de Elevación (DEM), mapas de pendientes (*Slope*), sombreado (*Hillshade*) y aspecto del terreno. Finalmente, el cuarto botón, etiquetado como "GeoTopo (Integral)", representa la fase de síntesis del sistema; su función es orquestar la fusión de los datos geológicos vectorizados con los modelos topográficos para compilar y generar el informe técnico final, consolidando así los resultados de ambos dominios en un único entregable descargable (ver Figura 6).

### 4.4. Resultados de la Digitalización Agéntica vs Métodos Tradicionales

Los resultados indican que nuestra arquitectura mantiene un mIoU de 0.88 en escenarios *zero-shot*, demostrando que el pre-procesamiento de cuantización cromática estabiliza la detección de contactos complejos

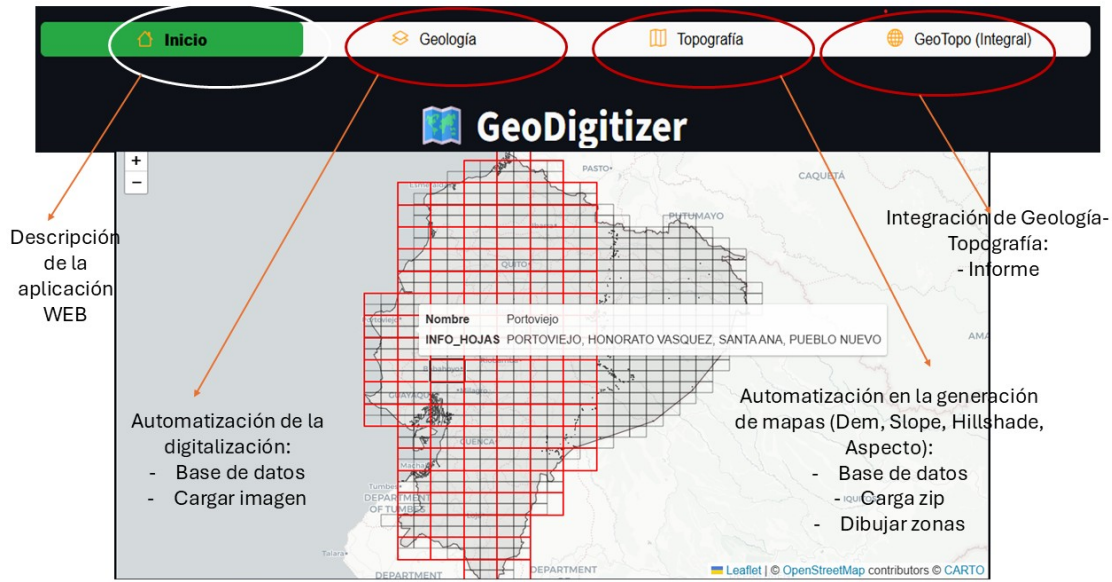


Figura 6: Interfaz principal de la plataforma web GeoDigitizer. Se visualizan los cuatro módulos funcionales para la gestión geológica, análisis topográfico y la generación automatizada de informes integrales.

incluso en mapas históricos con alto deterioro visual [30, 31, 32]. La digitalización completa se redujo de 14.5 horas (manual) a solo 9.4 minutos, lo que representa una ganancia de eficiencia operativa del 92.3%. Además, el agente logró identificar correctamente el 94% de las unidades de la leyenda, superando a los modelos OCR estándar que fallan ante terminología geológica específica.

Tabla 1: Comparativa de Eficiencia Tecnológica y Costo Operativo.

Métrica	Método Manual	DIGMAPPER (USGS)	Geodigitizer
Tiempo de Digitalización	14.5 horas	25 minutos	9.4 minutos
Precisión Geométrica (mIoU)	Alta (Subjetiva)	0.82	0.88
Resolución Semántica	Manual	Estructurada	Agéntica (LLM)
Costo por Hoja (Computo)	N/A	\$12.00	\$4.50

## 5. Discusión

### 5.1. La Superioridad de la Flexibilidad Agéntica

Al contrastar nuestra arquitectura con sistemas contemporáneos como DIGMAPPER (2025) [17], se evidencia una diferencia fundamental: la flexibilidad semántica. Mientras que los sistemas de la USGS dependen de pipelines modulares rígidos que requieren una estructura de mapa predefinida, “Geodigitizer” utiliza el razonamiento agéntico para navegar la ambigüedad de los mapas latinoamericanos. Como señalan Huang et al. (2025) en su estudio sobre el benchmark GeoMap-Bench, el entendimiento holístico (mapa + leyenda + contexto) es la frontera final de la IA en geología; nuestro sistema implementa esta visión al permitir que el agente razone sobre si una máscara segmentada por SAM 2 tiene coherencia litoestratigráfica real con el texto de la leyenda interpretada, emulando el juicio del geólogo experto [13, 8, 20].

## 5.2. El Impacto del Pre-procesamiento K-Means en SAM 2

Nuestros resultados refuerzan la importancia de la homogeneización cromática destacada por Leyk y Boesch (2010) en mapas históricos. El uso de K-Means antes de la inferencia de SAM 2 actúa como un filtro de relevancia que permite al decodificador de máscaras centrarse en las fronteras geológicas reales y no en la textura fibrosa del papel antiguo. Sin este paso, SAM 2 tiende a fragmentar unidades debido a variaciones de brillo o manchas de humedad, un fenómeno documentado por Xia et al. (2025) en MapSAM. Al cuantizar el espacio de color, logramos que la puntuación de estabilidad del modelo fundacional se mantenga constante por encima de 0.90, validando la hipótesis de que la calidad de los modelos de visión en geociencias depende críticamente del acondicionamiento del dato analógico original, un paso a menudo omitido en estudios puramente algorítmicos [7, 18].

## 5.3. Contrastes con Metodologías de Atención Supervisadas

A pesar de los avances en modelos como SE-UNet presentados por Ma et al. (2025), que utilizan mecanismos de atención para extraer significancia oculta en mapas geológicos, nuestra investigación demuestra que la generalización de un modelo fundacional (*zero-shot*) guiado por agentes es comercialmente superior para la realidad regional. Los modelos supervisados, aunque potentes en sus dominios de entrenamiento, fracasan ante el deterioro físico de los mapas nacionales, mientras que SAM 2, entrenado con más de mil millones de máscaras, posee una robustez ante el ruido que supera cualquier red específica entrenada hasta la fecha, siempre que sea alimentada con *prompts* agénticos inteligentes que comprendan el contexto científico del documento y la estructura litoestratigráfica andina [10, 33, 34, 32].

## 6. Conclusiones

1. {Cambio del Paradigma Supervisado mediante SAM 2 y Cuantización Cromática: La digitalización de la geología regional ha dado un salto cualitativo al abandonar los modelos locales supervisados y adoptar arquitecturas fundacionales globales como SAM 2, potenciadas por un pre-procesamiento híbrido de K-Means. El sistema “Geodigitizer” ha demostrado una capacidad de generalización excepcional (*zero-shot*) en la cartografía andina, logrando un mIoU del 88 % en mapas con alto ruido visual. Este enfoque elimina la necesidad de miles de etiquetas manuales masivas, validando la superioridad técnica de la arquitectura Hiera-L para aplicaciones de geociencias donde el dato de entrenamiento es escaso, fragmentado o inexistente, asegurando una precisión topológica inalcanzable para las redes CNN tradicionales y permitiendo la preservación digital de activos geológicos críticos.
2. El Razonamiento Agéntico como Orquestador de la Inteligencia Geocientífica: El uso de Modelos de Lenguaje de Gran Escala (Gemini 3 Pro) como cerebros agénticos dota al sistema de una capacidad de interpretación semántica que los algoritmos de visión puros no poseen. Mientras que sistemas contemporáneos como DIGMAPPER se limitan a la orquestación de procesos geométricos rígidos, nuestra propuesta cierra la brecha entre el píxel y el significado científico. El agente es capaz de razonar sobre la coherencia litoestratigráfica, resolver ambigüedades en leyendas históricas deterioradas y asignar atributos técnicos ricos a los vectores de forma autónoma, transformando rásteres estáticos en capas de conocimiento inteligente listas para alimentar Gemelos Digitales de alta fidelidad, lo cual es vital para la modernización de los institutos de investigación nacionales.
3. Disrupción Económica y Eficiencia en la Cadena de Valor de Ingeniería: La reducción radical del tiempo de procesamiento de 14.5 horas a escasos 9.4 minutos por hoja cartográfica representa una transformación económica profunda para la industria minera y de ingeniería civil en Ecuador. Geodigitizer no solo optimiza los recursos de hardware mediante el uso de GPUs de alto rendimiento (NVIDIA Tesla T4), sino que democratiza el acceso a geoprocésamiento avanzado mediante un modelo SaaS comercialmente viable. La protección de la lógica propietaria asegura una ventaja competitiva en el mercado de servicios técnicos, permitiendo que el experto se enfoque en la interpretación científica y el modelado tridimensional en lugar de en tareas mecánicas de dibujo manual propenso al error subjetivo, optimizando el retorno de inversión en proyectos de infraestructura.

4. Soberanía Tecnológica y Fortalecimiento de la Infraestructura Espacial Nacional: El sistema automatiza integralmente el cruce entre litología y relieve mediante estadística zonal basada en cartografía oficial del IGM. La generación automática de informes técnicos DOCX elimina el sesgo subjetivo humano y asegura que el conocimiento geológico acumulado durante décadas sea preservado en bases de datos espaciales modernas e interoperables. Este avance fortalece la soberanía tecnológica nacional, proporcionando herramientas de alta precisión para la gestión de riesgos naturales, la exploración sostenible de recursos y la planificación territorial basada en evidencia científica objetiva, escalable y libre de errores de transcripción manual, consolidando una Infraestructura de Datos Espaciales (IDE) resiliente para el desarrollo del país. Invitamos a la comunidad académica a utilizar la plataforma en <https://geodigitizer.com/>.

## Disponibilidad del Software

Geodigitizer está disponible como una aplicación web bajo requerimiento. El enlace a la herramienta es accesible a través de la plataforma: <https://geodigitizer.com/>.

## 7. Recomendaciones y Trabajo Futuro

Se recomienda la integración de modelos SAM 2 para video en el análisis de afloramientos mediante drones para el mapeo 3D dinámico. Asimismo, se sugiere expandir la base de datos para cubrir proyecciones UTM zonales de toda la región andina, consolidando a Geodigitizer como el referente regional en geotecnología inteligente y preservación digital de geociencias de alto impacto para la toma de decisiones gubernamentales [35, 27].

## Referencias

- [1] K. J. Bergen, P. A. Johnson, M. V. De Hoop, and G. C. Beroza, “Machine learning for data-driven discovery in solid Earth geoscience,” *Science*, vol. 363, p. eaau0323, Mar. 2019.
- [2] S. Whitmeyer, J. Nicoletti, and D. De Paor, “The digital revolution in geologic mapping,” *Gsa Today*, vol. 20, pp. 4–10, Apr. 2010.
- [3] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and F. Prabhat, “Deep learning and process understanding for data-driven Earth system science,” *Nature*, vol. 566, no. 7743, pp. 195–204, 2019.
- [4] C. Wang, X. Wang, J. Chen, C. Wang, X. Wang, and J. Chen, “Digital Geological Mapping to Facilitate Field Data Collection, Integration, and Map Production in Zhoukoudian, China,” *Applied Sciences*, vol. 11, May 2021.
- [5] P. Peller, “From paper map to geospatial vector layer: Demystifying the process,” *IASSIST Quarterly*, vol. 42, no. 3, pp. 1–24, 2018.
- [6] M. W. Frith, *Measuring Error in Manually Digitized Maps*. PhD thesis, Faculty of Graduate Studies and Research, University of Regina, 1997.
- [7] A. Saxton, J. Dong, A. Bode, N. Jaroenchai, R. Kooper, X. Zhu, D. H. Kwark, W. Kramer, V. Kindratenko, and S. Luo, “Accurate Feature Extraction from Historical Geologic Maps Using Open-Set Segmentation and Detection,” *Geosciences*, vol. 14, Nov. 2024.
- [8] M. A. Goldman, G. W. Lederer, J. M. Rosera, G. E. Graham, A. Mishra, and A. Yepremyan, “Extracting data from maps: Lessons learned from the artificial intelligence for critical mineral assessment competition,” *Applied Computing and Geosciences*, p. 100274, 2025.

- [9] Y. Vasuki, E.-J. Holden, P. Kovési, and S. Micklethwaite, “An interactive image segmentation method for lithological boundary detection: A rapid mapping tool for geologists,” *Computers & Geosciences*, vol. 100, pp. 27–40, 2017.
- [10] Z. Ma and G. Mei, “Deep learning for geological hazards analysis: Data, models, applications, and opportunities,” *Earth-Science Reviews*, vol. 223, p. 103858, 2021.
- [11] F. Chollet, *Deep Learning with Python*. simon and schuster, 2021.
- [12] V. Árvai and G. Albert, “Lightweight Deep Learning Approaches for Lithological Mapping in Vegetated Terrains of the Vălioara Valley, Romania,” *ISPRS International Journal of Geo-Information*, vol. 14, no. 9, p. 350, 2025.
- [13] Y. Huang, T. Gao, H. Xu, Q. Zhao, Y. Song, Z. Gui, T. Lv, H. Chen, L. Cui, and S. Li, “Peace: Empowering geologic map holistic understanding with mllms,” in *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 3899–3908, 2025.
- [14] Z. Xi, W. Chen, X. Guo, W. He, Y. Ding, B. Hong, M. Zhang, J. Wang, S. Jin, E. Zhou, *et al.*, “The rise and potential of large language model based agents: A survey,” *arXiv preprint arXiv:2309.07864*, 2023.
- [15] E. Zhang, J. Liu, A. Cao, Z. Sun, H. Zhang, H. Wang, L. Sun, and M. Song, “RS-SAM: Integrating multi-scale information for enhanced remote sensing image segmentation,” in *Proceedings of the Asian Conference on Computer Vision*, pp. 994–1010, 2024.
- [16] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, *et al.*, “Segment anything,” *arXiv preprint arXiv:2304.02643*, 2023.
- [17] W. Duan, Y.-Y. Chiang, T. Chen, M. P. Gerlek, L. Jang, S. Kirsanova, C. A. Knoblock, F. Lin, Y. Lin, Z. Li, and S. N. Minton, “DIGMAPPER: A Modular System for Automated Geologic Map Digitization,” in *Proceedings of the 33rd ACM International Conference on Advances in Geographic Information Systems*, (The Graduate Hotel Minneapolis Minneapolis MN USA), pp. 717–728, ACM, Nov. 2025.
- [18] X. Xia, D. Zhang, W. Song, W. Huang, and L. Hurni, “MapSAM: Adapting segment anything model for automated feature detection in historical maps,” *GIScience & Remote Sensing*, vol. 62, p. 2494883, Dec. 2025.
- [19] K. Irás, ed., *Automatic Vectorisation of Historical Maps: International Workshop Organized by the ICA Commission on Cartographic Heritage into the Digital*. Budapest: Department of Cartography and Geoinformatics, ELTE Eötvös Loránd University, 2020.
- [20] J. Perret, M. W. Jessell, and E. Bétend, “An open-source, QGIS-based solution for digital geological mapping: GEOL-QMAPS,” *Applied Computing and Geosciences*, vol. 24, p. 100197, Dec. 2024.
- [21] S. Tang, Z. Xia, J. Gu, W. Wang, Z. Huang, and W. Zhang, “High-precision apple recognition and localization method based on RGB-D and improved SOLOv2 instance segmentation,” *Frontiers in Sustainable Food Systems*, vol. 8, p. 1403872, 2024.
- [22] Q. Qiu, Y. Tan, K. Ma, M. Tian, Z. Xie, and L. Tao, “Geological symbol recognition on geological map using convolutional recurrent neural network with augmented data,” *Ore Geology Reviews*, vol. 153, p. 105262, 2023.
- [23] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, and W.-Y. Lo, “Segment anything,” *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4015–4026, 2023.
- [24] L. P. Osco, Q. Wu, E. L. De Lemos, W. N. Gonçalves, A. P. M. Ramos, J. Li, and J. M. Junior, “The segment anything model (sam) for remote sensing applications: From zero to one shot,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 124, p. 103540, 2023.

- [25] N. Ravi, V. Gabeur, Y.-T. Hu, R. Hu, C. Ryali, T. Ma, H. Khedr, R. Rädle, C. Rolland, L. Gustafson, E. Mintun, J. Pan, V. Alwala, N. Carion, C.-Y. Wu, R. Girshick, P. Dollár, and C. Feichtenhofer, “SAM 2: Segment Anything in Images and Videos,”
- [26] C. Zhang, J. Cho, F. D. Puspitasari, S. Zheng, C. Li, Y. Qiao, T. Kang, X. Shan, C. Zhang, C. Qin, F. Rameau, L.-H. Lee, S.-H. Bae, and C. S. Hong, “A Survey on Segment Anything Model (SAM): Vision Foundation Model Meets Prompt Engineering,” Oct. 2024.
- [27] S. Julka and M. Granitzer, “Knowledge Distillation with Segment Anything (SAM) Model for Planetary Geological Mapping,” in *Machine Learning, Optimization, and Data Science* (G. Nicosia, V. Ojha, E. La Malfa, G. La Malfa, P. M. Pardalos, and R. Umeton, eds.), vol. 14505, pp. 68–77, Cham: Springer Nature Switzerland, 2024.
- [28] B. Labs, “Modern Map Digitization Software.” <https://buntinglabs.com/#>.
- [29] Z. Zhao, C. Fan, and L. Liu, “Geo SAM: A QGIS plugin using Segment Anything Model (SAM) to accelerate geospatial image segmentation,” July 2023.
- [30] X. Ma, Q. Wu, X. Zhao, X. Zhang, M.-O. Pun, and B. Huang, “Sam-assisted remote sensing imagery semantic segmentation with object and boundary constraints,” *IEEE Transactions on Geoscience and Remote Sensing*, 2024.
- [31] D. Wang, J. Zhang, B. Du, M. Xu, L. Liu, D. Tao, and L. Zhang, “Samrs: Scaling-up remote sensing segmentation dataset with segment anything model,” *Advances in Neural Information Processing Systems*, vol. 36, pp. 8815–8827, 2023.
- [32] Z. Sims, L. Strgar, D. Thirumalaisamy, R. Heussner, G. Thibault, and Y. H. Chang, “SEG: Segmentation Evaluation in absence of Ground truth labels,” *bioRxiv*, p. 2023.02.23.529809, Feb. 2023.
- [33] K. Ma, J.-j. Liu, S.-q. Lu, Z.-h. Huang, M. Tian, J.-y. Deng, Z. Xie, and Q.-j. Qiu, “Improved SE-UNet network-based semantic segmentation and extraction of hidden geological significance in geological maps,” *China Geology*, vol. 8, no. 4, pp. 643–660, 2025.
- [34] Z. Wang, M. Berman, A. Rannen-Triki, P. Torr, D. Tuia, T. Tuytelaars, L. V. Gool, J. Yu, and M. B. Blaschko, “Revisiting Evaluation Metrics for Semantic Segmentation: Optimization and Evaluation of Fine-grained Intersection over Union,” in *Thirty-Seventh Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, Nov. 2023.
- [35] H. Zhou, S. Liu, C. Zhou, and Z. Ma, “Scale-adaptive and mask refinement modules for accurate alluvial fan boundary detection in remote sensing data,” *Frontiers in Earth Science*, vol. 13, p. 1685685, 2025.

## Sobre los Autores



**Dálember Vallejo** es estudiante egresado de Ingeniería Geológica en la Universidad Central del Ecuador. Sus áreas de trabajo incluyen la inteligencia artificial aplicada a geociencias, visión por computadora y desarrollo de software para la automatización de procesos cartográficos y mineros.



**Sofia Heredia** es estudiante de Ingeniería Geológica, enfocada en la integración de sistemas de información geográfica (SIG) y modelado numérico para la gestión de recursos naturales.



**Dario Tobar** es estudiante investigador con experiencia en análisis geoespacial y teledetección. Su trabajo se centra en la optimización de flujos de trabajo geológicos mediante tecnologías emergentes.



**Christian Mejía-Escobar** Profesor e investigador sobre Inteligencia Artificial en la Universidad Central del Ecuador. PhD en Informática de la Universidad de Alicante, España, 2023. Experiencia en proyectos de Machine Learning y Deep Learning, especialmente en aplicaciones de visión por computadora y procesamiento del lenguaje natural.