


GEOSEISMICAI: INTERPRETACIÓN AUTOMÁTICA DE SISMOFACIES MEDIANTE IA AGÉNTICA

A PREPRINT

Ronny López 

Silvana Hermosa 

Génesis Campos 

Sonia Gualán 

Christian Mejia-Escobar 

Facultad de Ingeniería en Geología, Minas, Petróleos y Ambiental (FIGEMPA)
Universidad Central del Ecuador
Quito, Ecuador

{ralopezt, smhermosa, gdcampos, sdgualan, cimejia}@uce.edu.ec

ABSTRACT

La interpretación sísmica es una tarea fundamental para comprender la geometría del subsuelo y su arquitectura estratigráfica, especialmente en las etapas iniciales de la exploración geológica. Tradicionalmente, este proceso se realiza de forma manual y depende en gran medida de la experiencia y el criterio subjetivo del intérprete, lo que implica una alta inversión de tiempo y posibles inconsistencias en la identificación de patrones complejos de sismofacies. En este trabajo se propone un sistema automatizado para la identificación y delimitación asistida de sismofacies en imágenes sísmicas bidimensionales, integrando técnicas de visión por computadora y aprendizaje profundo dentro de un flujo de trabajo reproducible. El núcleo del sistema se basa en la arquitectura DeepLabv3, que realiza extracción de características multiescala mediante convoluciones dilatadas e incorpora el módulo Atrous Spatial Pyramid Pooling (ASPP) para capturar información contextual a diferentes escalas. El modelo fue entrenado mediante transfer learning, alcanzando una accuracy de 0.891 en entrenamiento y 0.863 en validación. La evaluación sobre el conjunto de prueba evidenció un equilibrio adecuado entre precisión (Precision) y sensibilidad (Recall) por clase, demostrando una buena capacidad de generalización en la clasificación automática de las distintas sismofacies. Adicionalmente, el sistema integra un modelo multimodal basado en Transformers de OpenAI, capaz de procesar conjuntamente información visual y textual para generar interpretaciones geológicas automatizadas en lenguaje natural a partir de las sismofacies previamente identificadas. Todo el flujo de trabajo se orquesta mediante una arquitectura multiagente implementada en n8n y desplegada en una aplicación web interactiva que permite procesar imágenes sísmicas, visualizar resultados en tiempo real y generar reportes técnicos en formato PDF. GeoSeismicAI constituye una herramienta de apoyo para el análisis sísmico inicial, contribuyendo a mejorar la productividad, reproducibilidad y estandarización del proceso interpretativo sin sustituir el criterio experto del geólogo.

Palabras clave: Interpretación sísmica, sismofacies, segmentación semántica, aprendizaje profundo, DeepLabv3, visión por computadora, lenguaje natural

1. Introducción

Las sismofacies se definen como unidades sísmicas compuestas por grupos de reflexiones cuyas características como configuración geométrica, continuidad, amplitud y frecuencia difieren de los grupos adyacentes. Esto representa la base de la interpretación sísmica, ya que, permiten tener una visión general e inicial de la línea sísmica y finalmente inferir variaciones litológicas y ambientes deposicionales que hacen posible reconstruir su historia geológica. Una línea sísmica es una sección transversal que representa la respuesta acústica de las interfaces del subsuelo ante la propagación de ondas elásticas [1]. El proceso general de la prospección sísmica se divide en

tres etapas críticas: adquisición de datos, procesamiento y, finalmente, la interpretación [2]. La interpretación, como fase culminante, consiste en transformar los datos procesados en un modelo geológico coherente [3]. Este proceso implica la descripción detallada de reflectores, donde el intérprete identifica la continuidad de horizontes, analiza configuraciones geométricas, y evalúa las variaciones de amplitud y frecuencia que definen las sismofacies [4]. Además, la interpretación tradicional requiere de la integración de datos de pozos para validar las secuencias estratigráficas identificadas [5].

La interpretación sísmica constituye una herramienta importante para comprender la estructura, geometría y composición del subsuelo, donde las líneas sísmicas son necesarias en la caracterización de acuíferos, minería, exploración hidrocarburífera e ingeniería civil [6]. La metodología típica se basa en la identificación de discontinuidades estructurales y cambios en los patrones de reflexión que sugieren variaciones litológicas [7]. No obstante, este enfoque tradicional demanda una alta inversión de tiempo y presenta una subjetividad inherente al criterio del especialista [8]. Para mitigar esto, se han desarrollado técnicas de atributos sísmicos que ayudan a resaltar rasgos físicos no evidentes a simple vista [9].

En los últimos años, la aplicación de inteligencia artificial (IA), específicamente mediante redes neuronales convolucionales (CNN), ha revolucionado este campo al permitir la extracción automática de características espaciales [10]. Trabajos previos han demostrado la eficacia de estas tecnologías con el uso de datos sísmicos; por ejemplo, Araya-Polo et al. [11] implementaron la detección automática de fallas sin necesidad de procesamiento convencional, mientras que Huang et al. [12] desarrollaron plataformas escalables para identificar rasgos geológicos complejos. Otros estudios han aplicado CNN para la clasificación de cuerpos salinos [13, 14, 15]. En años recientes, el aprendizaje profundo ha impulsado el desarrollo de arquitecturas tipo U-Net y modelos encoder-decoder para la segmentación automática de facies sísmicas, con variantes especializadas y revisiones que consolidan su uso en interpretación sísmica moderna [16, 17, 18]. No obstante, la interpretación de sismofacies no se ha automatizado para imágenes sino para datos sísmicos específicos.

Basado en estos avances, el presente trabajo propone un sistema para la detección y segmentación automática de sismofacies en imágenes 2D, integrando técnicas de preprocesamiento, segmentación semántica y visualización dentro de un flujo automatizado. El núcleo del sistema se fundamenta en el modelo DeepLabv3, una arquitectura de tipo codificador-decodificador que incorpora convoluciones dilatadas para capturar información contextual a múltiples escalas, capacidad crítica ante las variaciones locales y regionales de las sismofacies [19]. Previo al entrenamiento, las imágenes son sometidas a un flujo de preprocesamiento que incluye conversión a escala de grises y realce de contraste con CLAHE, optimizando la nitidez de los reflectores [20]. El entrenamiento se realiza bajo un enfoque supervisado utilizando un conjunto de datos organizado y anotado mediante la visión por computadora de la plataforma RoboFlow, proporcionando las máscaras de referencia necesarias para el aprendizaje de las clases.

Adicionalmente, el sistema no solo segmenta automáticamente sismofacies mediante DeepLabv3, sino que también emplea procesamiento de lenguaje natural (PLN) con modelos de lenguaje LLM para generar interpretaciones asistidas. Todo el sistema se integra dentro de una arquitectura de tipo agéntica, donde la inteligencia artificial actúa como un mecanismo de orquestación que coordina de forma secuencial y reproducible las etapas de análisis, emulando el flujo técnico de un especialista [21, 22].

Como resultado final, se implementa un sitio web desarrollado con GitHub y Streamlit, donde el sistema genera la segmentación automática y devuelve la imagen interpretada en formato JPG. Además, consolida la información en un reporte técnico en formato PDF que incluye la imagen original, las sismofacies clasificadas y la interpretación asistida.

Este enfoque busca apoyar al geólogo en las etapas iniciales de análisis, incrementando la productividad y proporcionando una interpretación preliminar estandarizada, sin sustituir el criterio experto final del intérprete.

2. Metodología

El presente trabajo tiene como objetivo la detección, segmentación e interpretación automática de sismofacies en imágenes sísmicas bidimensionales mediante técnicas de inteligencia artificial, integrando herramientas de visión por computadora, aprendizaje profundo y procesamiento del lenguaje natural dentro de un sistema multiagente. En esta sección se describe la metodología propuesta, la cual se estructura en una serie de etapas secuenciales que conforman el flujo completo de análisis y que se resumen esquemáticamente en la Figura 1.

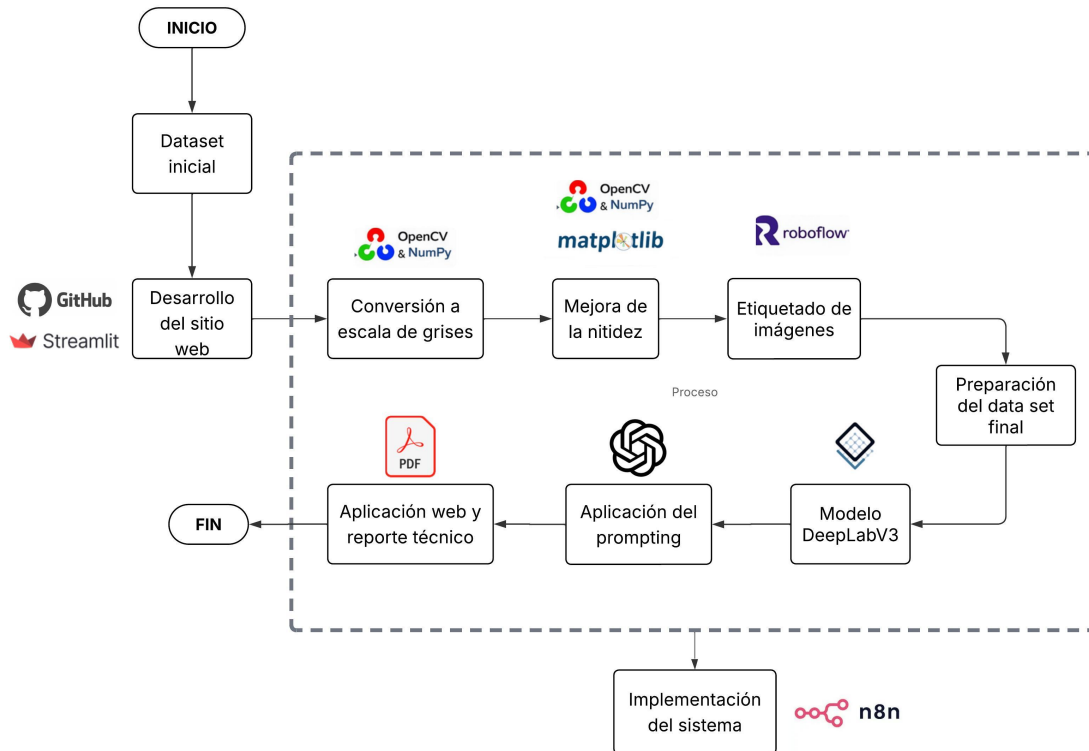


Figura 1: Flujograma general de las etapas principales de la metodología, incluyendo las plataformas y herramientas computacionales empleadas.

En primer lugar, se aborda la construcción de la materia prima necesaria para los modelos de aprendizaje automático, es decir, el conjunto de datos (dataset) de imágenes sísmicas. En segundo lugar, el dataset es preparado mediante un flujo de preprocesamiento orientado a mejorar la calidad visual de las imágenes y asegurar su consistencia. A continuación, las imágenes son sometidas a un proceso de etiquetado y anotación mediante herramientas de visión por computadora implementadas en la plataforma Roboflow, lo que permite delimitar y clasificar las distintas sismofacies presentes en cada sección sísmica de manera sistemática. Posteriormente, se entrena un modelo basado en redes neuronales convolucionales profundas (DeepLabv3) para la identificación automática de sismofacies. Una vez obtenidas las máscaras de segmentación, se incorpora un módulo de procesamiento del lenguaje natural que, a partir de los resultados generados por el modelo de segmentación, produce una interpretación geológica preliminar automatizada mediante un modelo de lenguaje multimodal. Finalmente, mediante el sistema multiagente implementado en n8n, se integra todo el flujo de trabajo y se desarrolla una aplicación web que permite la descarga de la imagen sísmica interpretada y la generación automática de un reporte técnico. A continuación, se detalla cada una de estas etapas.

2.1. Dataset inicial

El objetivo de esta etapa es recopilar un conjunto representativo de imágenes sísmicas que permita entrenar y evaluar el sistema propuesto. Las imágenes sísmicas fueron recopiladas a partir del libro *Atlas of Structural Geological Interpretation from Seismic Images* [4], una referencia ampliamente reconocida en estudios de interpretación sísmica. El dataset inicial está conformado por 200 imágenes sísmicas bidimensionales, incluyendo líneas interpretadas y no interpretadas, que muestran reflectores estratificados, variaciones de amplitud y geometrías asociadas a diferentes contextos estructurales y estratigráficos mostradas en la Figura 2.

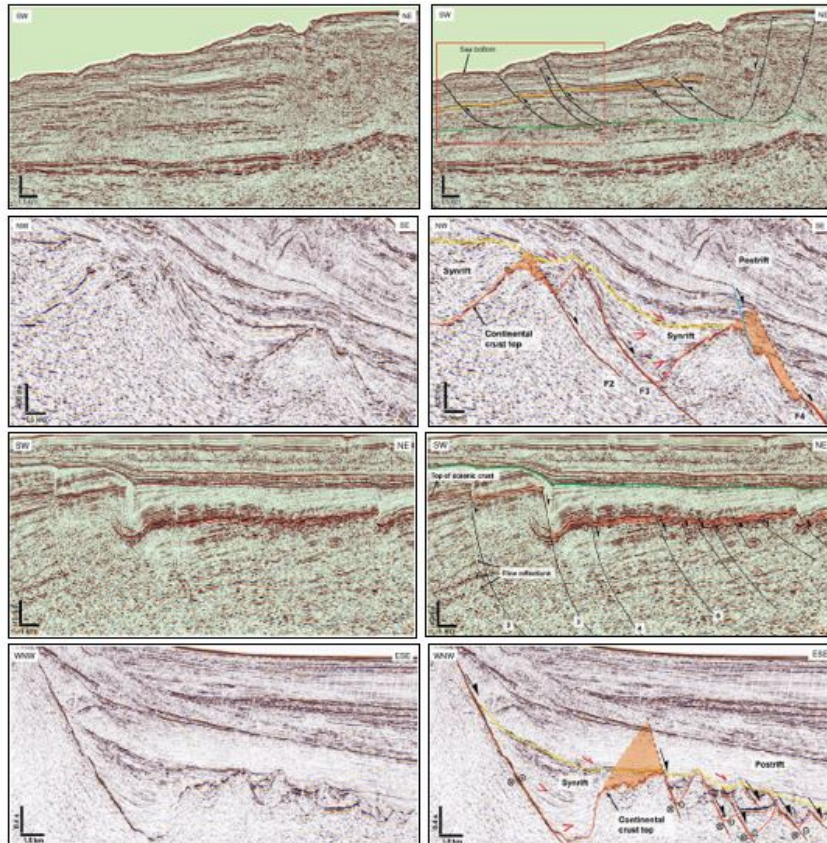


Figura 2: Ejemplos representativos de líneas sísmicas utilizadas en la construcción del dataset. La columna izquierda muestra secciones sísmicas originales no interpretadas, mientras que la columna derecha presenta las mismas secciones con interpretación geológica superpuesta (fallas, pliegues y contactos estratigráficos), utilizadas como referencia visual durante el proceso de etiquetado.

Los datos fueron almacenados inicialmente en un repositorio en Google Drive, organizados bajo una estructura jerárquica compuesta por dos carpetas principales: una destinada a las imágenes sísmicas originales (sin interpretación) y otra a las imágenes interpretadas según el libro de referencia. Esta organización permitió mantener una separación clara entre la información base y su correspondiente interpretación geológica.

Las imágenes fueron conservadas en formatos PNG y JPG, con resoluciones aproximadas de 4096×4096 píxeles y un tamaño promedio cercano a 1 MB por archivo, garantizando un nivel de detalle adecuado para el análisis de sismofacies.

La nomenclatura de los archivos siguió un esquema secuencial estructurado que permitió mantener la trazabilidad entre cada imagen original y su versión interpretada. Por ejemplo, la imagen original fue almacenada como *imagen_001.png*, mientras que su correspondiente interpretación fue nombrada como *interpretada_001.png*. El identificador numérico compartido (“001”) ase-

guró la correspondencia directa entre ambos archivos, facilitando su posterior procesamiento, etiquetado y validación dentro del flujo metodológico.

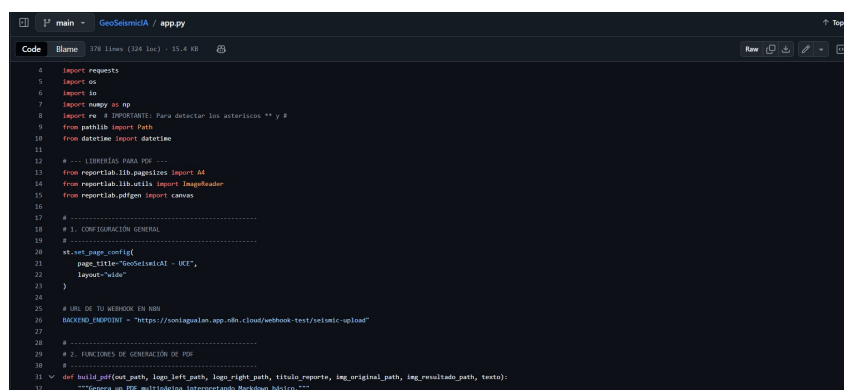
2.2. Desarrollo del sitio web

La implementación de GeoSeismicAI se realizó en lenguaje Python, utilizando Streamlit como framework principal para la generación de la aplicación web interactiva. Esta herramienta permitió desarrollar una interfaz dinámica capaz de gestionar la carga de imágenes sísmicas, ejecutar el modelo de segmentación y visualizar los resultados de manera inmediata, facilitando la interacción directa del usuario con el sistema.

El código fuente del proyecto fue desarrollado y versionado en GitHub¹, plataforma empleada como repositorio central para la construcción y mantenimiento de la aplicación web. La vinculación del repositorio con Streamlit permitió el despliegue automático de la aplicación, donde las dependencias definidas en el archivo requirements.txt se instalan de manera automática y el script principal (app.py) se ejecuta para generar la interfaz accesible desde cualquier navegador. Para el procesamiento de imágenes sísmicas se emplearon las bibliotecas Pillow (PIL) y NumPy, que permiten la lectura, conversión y manipulación eficiente de arreglos matriciales, así como la generación de máscaras de segmentación y superposiciones visuales entre la imagen original y los resultados del modelo. La librería Requests se utilizó para gestionar la comunicación entre el frontend y servicios externos mediante peticiones HTTP, mientras que los módulos Base64, io y Pathlib facilitaron la codificación de archivos, la manipulación de flujos en memoria y la gestión de rutas y archivos temporales dentro de la aplicación.

Por otro lado, para la generación automática de reportes técnicos en formato PDF se implementó la librería ReportLab, permitiendo integrar imágenes segmentadas, resultados e interpretación preliminar en un documento estructurado descargable directamente desde la interfaz web. Este conjunto de herramientas consolida un entorno web funcional, reproducible y escalable, cuya actualización puede realizarse directamente desde el repositorio en GitHub.

En la Figura 3 se presenta un fragmento representativo del código fuente, donde se observa la importación e integración de las principales librerías utilizadas, evidenciando la estructura que sustenta la construcción de la interfaz web.



```

Code | blame | 178 lines (124 loc) | 15.4 KB | Raw | Top
4 import requests
5 import os
6 import io
7 import numpy as np
8 import cv2 # IMPORTANTE: Para detectar los asteriscos ** y #
9 from pathlib import Path
10 from datetime import datetime
11
12 # --- LIBRERÍAS PARA PDF ---
13 from reportlab.lib.pagesizes import A4
14 from reportlab.lib.styles import ImageStyle
15 from reportlab.pdfgen import canvas
16
17 # -----
18 # 1. CONFIGURACIÓN GENERAL
19 # -----
20 st.set_page_config(
21     page_title="GeoSeismicAI - VEC",
22     layout="wide"
23 )
24
25 # URL DE TU WEBOOK EN RUM
26 BACKEND_ENDPOINT = "https://sonisgalan.app.robocloud/webhook-test/seismic-upload"
27
28 # -----
29 # 2. FUNCIONES DE GENERACIÓN DE PDF
30 # -----
31 def handle_pdf(upload_path, img_in_path, img_out_path, titolo_reporte, img_original_path, img_resultado_path, texto):
32     """Genera un PDF multipágina interpretando seismas sísmico."""

```

Figura 3: Fragmento del código fuente desarrollado en GitHub para la implementación del sitio web y la gestión de las librerías empleadas en la aplicación.

Antes del entrenamiento del modelo, el conjunto de imágenes sísmicas es sometido a un flujo de preprocesamiento estandarizado, diseñado para mejorar la calidad visual de las imágenes y garantizar la homogeneidad de los datos de entrada, tal como se indica en la Figura 4. Como etapa final de este proceso, las imágenes preprocesadas son sometidas a un procedimiento de etiquetado y anotación mediante herramientas de visión por computadora lo que permite definir

¹<https://github.com>

de manera sistemática las distintas sismofacies presentes y generar las máscaras necesarias para el entrenamiento supervisado del modelo.

2.2.1. Conversión a escala de grises

Las imágenes sísmicas son convertidas a escala de grises cuando se encuentran en formatos multicanal. Esta conversión se realiza mediante scripts automatizados en Python² utilizando las librerías OpenCV³ y NumPy⁴. La reducción de la información cromática a una sola banda de intensidad permite enfatizar patrones sísmicos relevantes, tales como la continuidad de reflectores, contrastes de amplitud y texturas estructurales, además de reducir la carga computacional durante el entrenamiento del modelo.

2.2.2. Mejora de nitidez

Posteriormente, se aplica la técnica CLAHE (*Contrast Limited Adaptive Histogram Equalization*) para mejorar el contraste local de las imágenes sísmicas (Figura 4). A diferencia de la ecualización global, CLAHE opera sobre subregiones de la imagen y limita la amplificación del contraste, evitando la introducción de ruido. Desde el punto de vista geofísico, esta técnica permite realzar bordes, discontinuidades y variaciones sutiles de amplitud, facilitando la identificación de sismofacies sin alterar la geometría original de los reflectores.

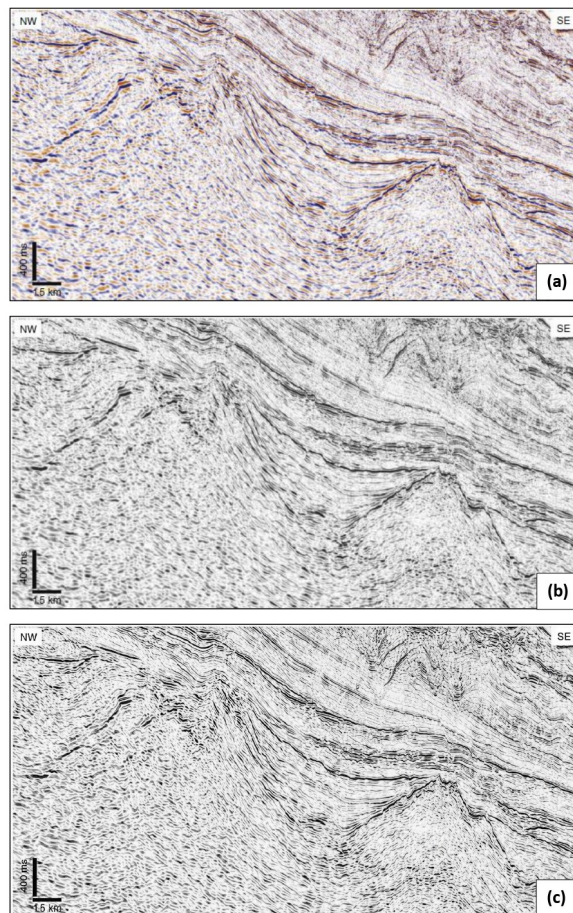


Figura 4: Ejemplo del flujo de preprocesamiento aplicado a una línea sísmica: (a) imagen original, (b) conversión a escala de grises, (c) mejora de contraste mediante CLAHE.

²<https://www.python.org>

³<https://opencv.org>

⁴<https://numpy.org>

2.2.3. Etiquetado de imágenes

Las imágenes sísmicas fueron sometidas a un proceso de etiquetado manual con el objetivo de generar las máscaras necesarias para el entrenamiento supervisado del modelo de Deep Learning. Este proceso constituye una de las etapas más críticas del flujo de trabajo, ya que la calidad del etiquetado influye directamente en la capacidad del modelo para aprender patrones geológicos representativos.

Las imágenes sísmicas utilizadas como entrada corresponden a secciones bidimensionales en escala de grises, en las cuales se observan reflectores estratificados, discontinuidades y variaciones de amplitud asociadas a diferentes contextos estructurales y estratigráficos.

El etiquetado se realizó mediante la plataforma Roboflow⁵, donde cada imagen fue clasificada mediante segmentación semántica para identificar distintas clases de sismofacies. La segmentación de las imágenes sísmicas se estructuró en 14 clases de sismofacies, definidas como combinaciones específicas entre la configuración interna de los reflectores y sus atributos sísmicos característicos. Cada clase representa una firma sísmica diferenciable dentro de la imagen y fue tratada como una categoría independiente en el proceso de anotación pixel a pixel. La nomenclatura adoptada sigue la estructura: Configuración-Amplitud-Frecuencia-Continuidad donde la configuración interna puede corresponder a patrones paralelos, subparalelos, caóticos o paralelos contorsionados, mientras que los atributos del reflector incluyen variaciones de amplitud alta o baja (AA, AB), frecuencia alta o baja (FA, FB) y continuidad o discontinuidad (C, D).

Estas combinaciones permiten capturar de manera integrada tanto la geometría interna del patrón reflectivo como sus propiedades dinámicas asociadas, generando clases compuestas que reflejan diferencias texturales y estructurales dentro de la sección sísmica. Cada una de las 14 clases fue delimitada manualmente durante el proceso de anotación, garantizando coherencia interpretativa y consistencia en la segmentación

En la Figura 5 se presenta un ejemplo del resultado del proceso de etiquetado, donde cada sismofacies es representada mediante un color distinto. Estas máscaras constituyen la verdad de referencia (*ground truth*) utilizada durante el entrenamiento del modelo de segmentación automática.

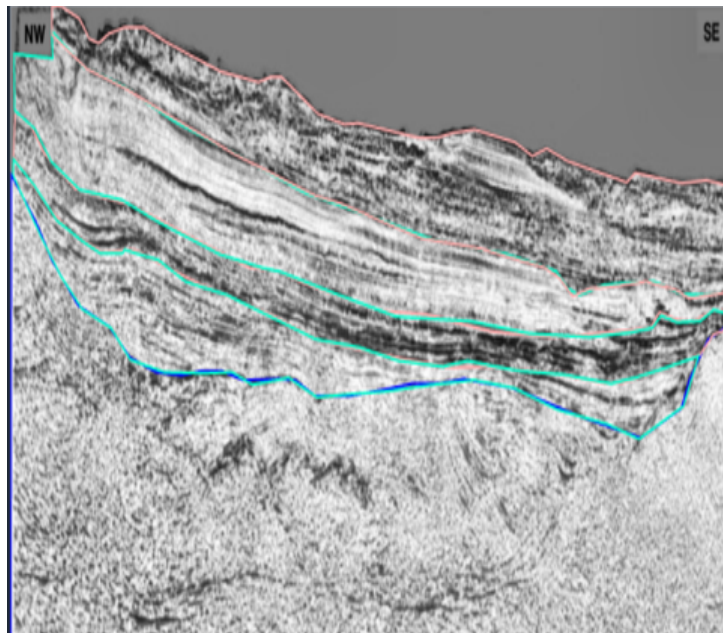


Figura 5: Ejemplo de sección sísmica con máscara de sismofacies asociada, donde cada color representa una clase diferente.

⁵<https://roboflow.com>

2.3. Preparación del dataset final

Con el fin de mejorar la capacidad de generalización del modelo, se aplicaron técnicas de aumento de datos (data augmentation), incluyendo rotaciones, escalamiento y transformaciones geométricas controladas. Como resultado, el conjunto inicial de 200 imágenes fue ampliado hasta un total de 340 imágenes anotadas. Posteriormente, el dataset fue organizado y dividido en tres subconjuntos independientes: entrenamiento (train), validación (val) y prueba (test), siguiendo una proporción del 80 %, 10 % y 10 %, respectivamente. Este proceso generó dos carpetas principales: una destinada a las imágenes originales y otra a las máscaras asociadas, manteniendo la correspondencia directa entre cada imagen y su respectiva anotación. La Tabla 1 resume la distribución final del conjunto de datos:

Tabla 1: Distribución y cantidad de imágenes para el dataset

Conjunto	Número de imágenes	Número de máscaras	Porcentaje
Train	272	272	80 %
Val	34	34	10 %
Test	34	34	10 %
Total	340	340	100 %

Cada imagen posee su correspondiente máscara de referencia (ground truth), generada mediante etiquetado manual en la plataforma Roboflow. Finalmente, el dataset final fue descargado en formato JSON COCO Segmentation, correspondiente a un esquema de anotación para clasificación pixel a pixel.

2.4. Modelo DeepLabV3

El algoritmo DeepLabv3, basado en redes neuronales convolucionales profundas, fue introducido por Chen et al. [19] para abordar el problema de la segmentación semántica de imágenes. Su propuesta se centra en mejorar la comprensión del contexto multiescala sin sacrificar la resolución espacial, mediante el uso de convoluciones atrous o dilatadas. A diferencia de enfoques tradicionales que dependen de arquitecturas encoder-decoder complejas o de múltiples etapas de procesamiento, DeepLabv3 analiza la imagen de entrada en una sola pasada convolucional profunda, incorporando información contextual a diferentes escalas de manera eficiente. Esta característica le permite alcanzar un alto desempeño en tareas de segmentación, manteniendo un equilibrio entre precisión semántica y costo computacional [19, 32]. Una de las principales fortalezas de DeepLabv3 es el módulo Atrous Spatial Pyramid Pooling (ASPP), el cual aplica convoluciones con distintas tasas de dilatación en paralelo, permitiendo capturar patrones locales y globales simultáneamente. Gracias a este diseño, el modelo ha demostrado una gran capacidad para segmentar estructuras complejas y continuas, lo que ha favorecido su adopción en diversas aplicaciones de visión por computadora incluyendo las geociencias. En este trabajo se emplea DeepLabv3 con backbone ResNet-101 preentrenado, siguiendo el esquema de transfer learning. La arquitectura se compone de tres bloques principales como se muestra en la Figura 6.

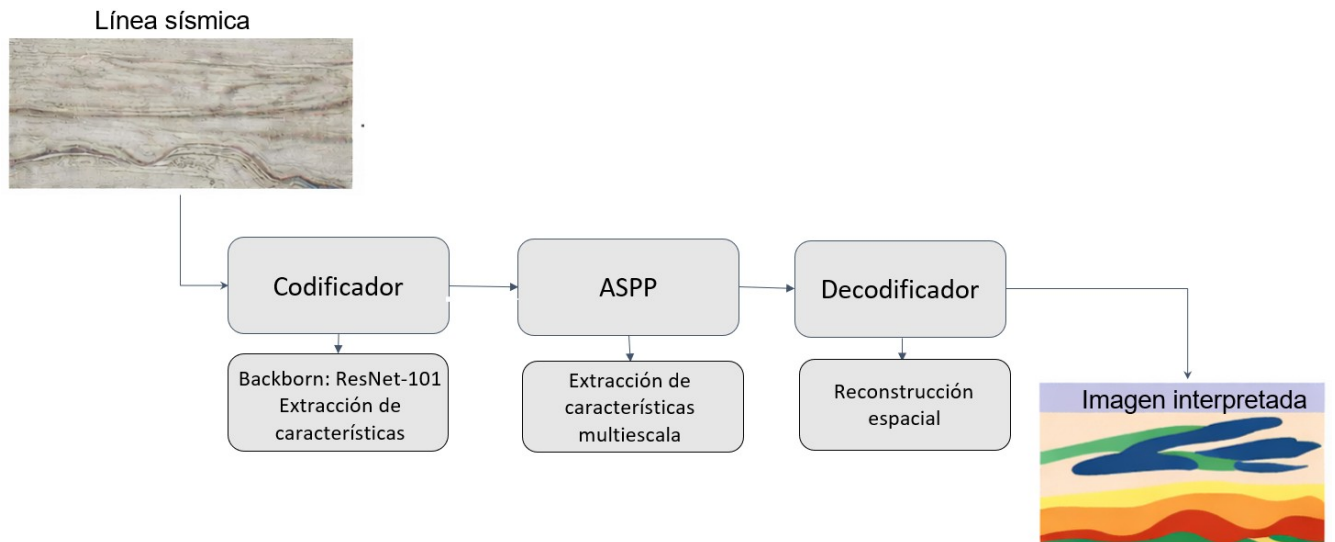


Figura 6: Arquitectura general del modelo DeepLabv3 empleado para la segmentación automática de sismofacies.

- Backbone convolucional (ResNet-101).** El backbone actúa como extractor jerárquico de características (CNN), aprendiendo representaciones espaciales y semánticas a diferentes escalas. ResNet-101 introduce conexiones residuales que facilitan el entrenamiento de redes profundas y mitigan el problema del *vanishing gradient*.
- Módulo ASPP (Atrous Spatial Pyramid Pooling).** El núcleo de DeepLabv3 es el módulo ASPP, el cual aplica convoluciones *atrous* (dilatadas) con diferentes tasas de dilatación en paralelo. Esto permite capturar información contextual a múltiples escalas sin perder resolución espacial. Los resultados se concatenan y se proyectan a un espacio de características común, proporcionando un equilibrio entre detalle local y contexto global (Figura 8).
- Decodificador.** En la implementación utilizada, el modelo DeepLabv3 no incorpora un decodificador explícito en el sentido clásico *encoder-decoder*. Tras la extracción de características mediante el backbone ResNet-101 y la agregación de contexto multiescala a través del módulo Atrous Spatial Pyramid Pooling (ASPP), el clasificador produce mapas de predicción a resolución reducida. Estos mapas son posteriormente reescalados al tamaño original de la imagen mediante interpolación bilineal, proceso que actúa como un mecanismo de decodificación implícita, permitiendo proyectar las predicciones semánticas al espacio original de la imagen sin realizar fusión con características de bajo nivel, tal como se describe en la arquitectura original de DeepLabv3 [19]. Posteriormente, el bloque decodificador combina la información de alto nivel proveniente del ASPP con características de bajo nivel extraídas en etapas tempranas del codificador. Esta fusión permite recuperar detalles espaciales finos, mejorando la precisión de los límites entre sismofacies. A través de operaciones de sobremuestreo y convoluciones adicionales, el decodificador reconstruye una máscara de segmentación con la misma resolución espacial que la imagen de entrada.

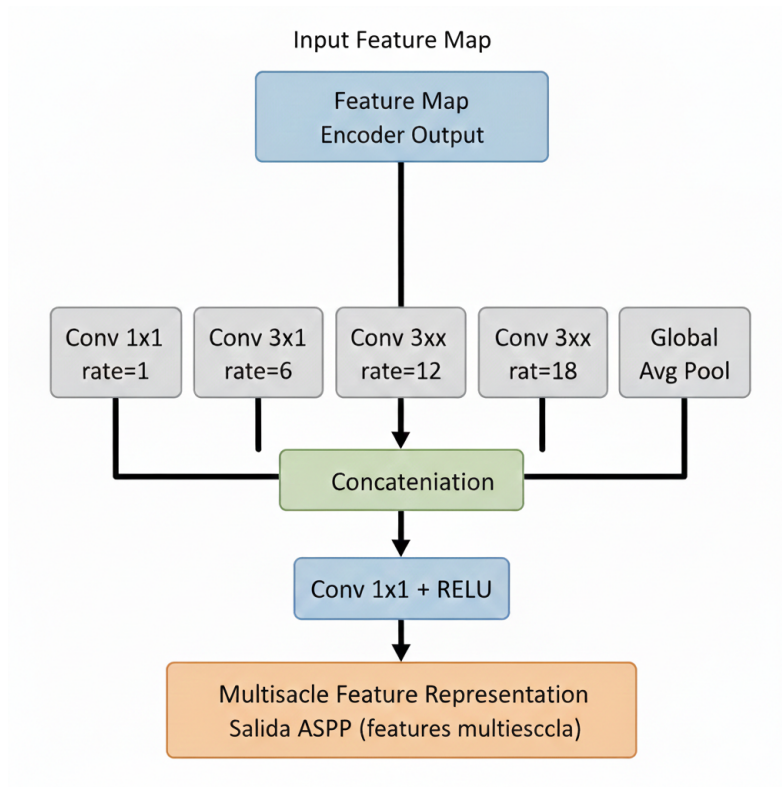


Figura 7: Representación esquemática del módulo Atrous Spatial Pyramid Pooling (ASPP) utilizado para capturar información multiescala.

Como resultado final, el modelo genera una máscara de segmentación donde cada píxel es clasificado en una de las clases de sismofacies definidas. En la Figura 8 se presenta un ejemplo representativo de la segmentación automática obtenida del modelo entrenado.

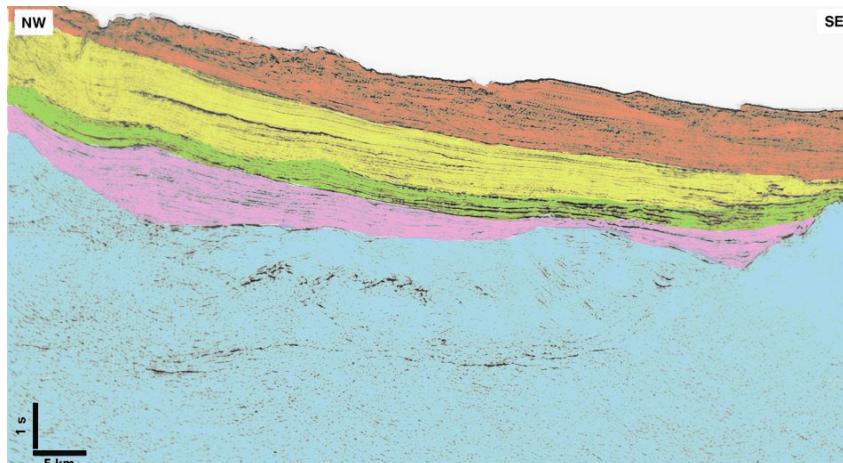


Figura 8: Ejemplo de segmentación automática de sismofacies generada por el modelo DeepLabv3. La unidad superior (naranja) muestra reflectores continuos y paralelos, indicando una sedimentación relativamente uniforme. Debajo, la unidad amarilla mantiene una disposición concordante, aunque con ligeras variaciones de espesor. La unidad rosada evidencia un cambio en la textura sísmica y en la continuidad lateral de los reflectores. La unidad inferior (celeste) presenta una textura menos organizada y reflectores poco definidos, asociándose al basamento.

2.5. Aplicación del prompting

Para la generación de interpretaciones geológicas a partir de las sismofacies segmentadas, se empleó el modelo de lenguaje multimodal GPT-4o, desarrollado por OpenAI ⁶. Mediante un prompt estructurado, el sistema recibe tanto la imagen segmentada como el texto descriptivo introducido por el usuario, generando posteriormente una interpretación geológica integral. Esta aproximación permite combinar la segmentación automática con capacidades avanzadas de razonamiento contextual y generación de lenguaje natural.

Con el fin de garantizar coherencia interpretativa y reducir la variabilidad en las respuestas generadas, el prompt fue diseñado siguiendo una estructura formal basada en principios de prompt engineering. Esta estructura organiza explícitamente los componentes fundamentales de la interacción con el modelo, incluyendo la definición del rol experto, el objetivo del análisis, el contexto de entrada, las instrucciones específicas, el formato de salida esperado y las restricciones semánticas.

En la Tabla 2 se presenta de manera resumida la configuración adoptada para el diseño del prompt dentro del sistema propuesto.

Tabla 2: Estructura del prompt utilizado para la interpretación geológica automatizada

Componente del Prompt	Descripción
Rol	Geólogo especialista en interpretación sísmica estratigráfica con experiencia en análisis de sismofacies.
Objetivo	Generar una interpretación geológica preliminar a partir de una imagen sísmica 2D segmentada en 14 clases de sismofacies.
Contexto	Imagen procesada mediante segmentación semántica pixel a pixel (DeepLabv3), donde cada clase representa una combinación entre configuración interna y atributos sísmicos (amplitud, frecuencia y continuidad).
Instrucciones	Analizar la imagen segmentada e identificar las unidades interpretadas de base a techo. Para cada unidad reconocida, describir de manera estructurada: (i) tipo de unidad, (ii) espesor aparente, (iii) interpretación tectónica, (iv) tipo de contacto y (v) ambiente sedimentario. La respuesta debe organizarse en unidades numeradas y concluir con una síntesis interpretativa que integre la evolución estructural y sedimentaria del sistema..
Salida	Interpretación organizada en descripción general por cada una de las sismofacies indicadas.
Restricciones	Evitar especulación no sustentada, no inventar información no observable y mantener tono académico objetivo.

Establecida la estructura del prompt y su lógica de interacción, se describe la arquitectura fundamental del modelo de OpenAI, la cual sustenta el procesamiento y la generación de las respuestas interpretativas. La base arquitectónica corresponde al Transformer, introducido por Vaswani (2017) [23] en el artículo “Attention Is All You Need”, el cual propuso un esquema encoder–decoder fundamentado en mecanismos de atención y autoatención. En su variante *decoder-only*, adoptada por la familia de OpenAI, el procesamiento comienza con la normali-

⁶<https://openai.com>

zación y tokenización del texto de entrada frecuentemente mediante sub-palabras, donde cada token es convertido en un índice numérico dentro de un vocabulario discreto. Estos índices son transformados posteriormente en vectores menos densos a través de una matriz de embeddings. A cada embedding léxico se le suma un embedding posicional, generalmente basado en funciones seno-coseno, que codifica la posición del token en la secuencia y permite preservar el orden secuencial. El resultado es una secuencia de vectores en un espacio latente continuo que atraviesa múltiples capas de masked self-attention, donde cada token atiende únicamente a los anteriores durante la generación autoregresiva del texto.

Para incorporar información visual, el sistema emplea un codificador basado en Vision Transformer (ViT), propuesto por Dosovitskiy (2020) [24]. En este esquema (Figura 10), la imagen correspondiente al mapa de sismofacies previamente generado por DeepLabV3 se divide en parches de tamaño fijo, los cuales son proyectados linealmente al espacio de embeddings visuales y enriquecidos con codificación posicional antes de pasar por capas de autoatención. Posteriormente, estos embeddings visuales se proyectan al mismo espacio latente que los embeddings lingüísticos y se integran como tokens dentro de la secuencia procesada por el decoder. De esta manera, el modelo aplica autoatención conjunta sobre información textual y visual, permitiendo que el proceso autoregresivo genere interpretaciones geológicas condicionadas tanto por los patrones espaciales segmentados como por el contexto semántico del prompt, integrando así análisis cuantitativo y razonamiento lingüístico para la generación de la interpretación geológica.

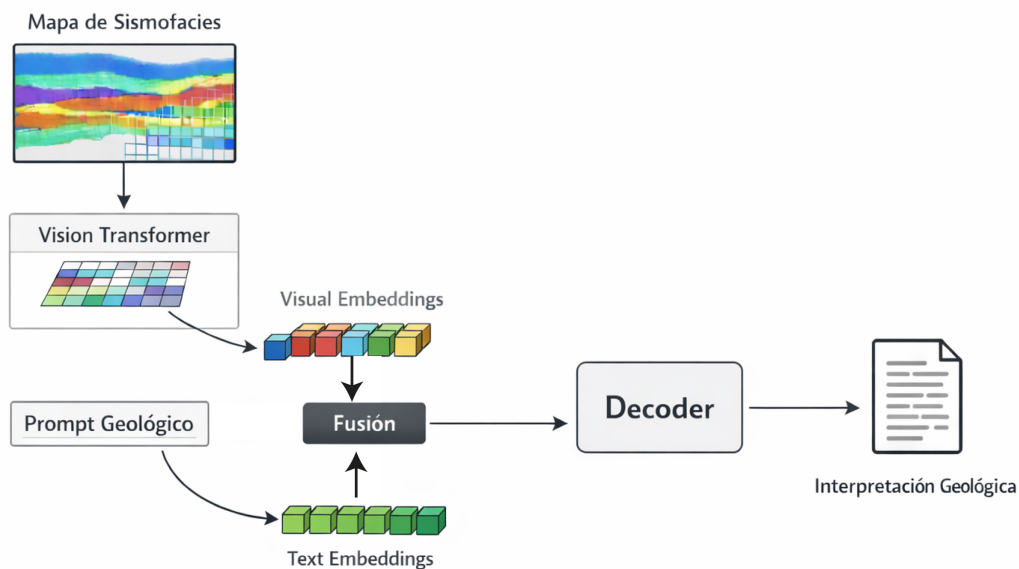


Figura 9: Esquema del modelo de transformer de Open AI.

2.6. Implementación del sistema

La arquitectura de GeoSeismicAI se fundamenta en un enfoque de Inteligencia Artificial Agéntica, en el cual cada componente del sistema actúa como un agente especializado capaz de ejecutar tareas específicas, validar condiciones de entrada y tomar decisiones que determinan la continui-

dad o interrupción del flujo de procesamiento. Este enfoque permita una ejecución autónoma, modula y adaptable de las etapas metodológicas del análisis sísmico. La orquestación del sistema se llevó a cabo mediante *n8n*⁷ (Figura 10), una plataforma de automatización de flujos de trabajo basada en nodos, que permite la integración modular de modelos de aprendizaje profundo y modelos de lenguaje grande (LLM). La interacción con el usuario final se gestiona a través de una interfaz web desarrollada en *Streamlit*⁸. La comunicación entre el frontend (Streamlit) y el orquestador (backend en *n8n*) se establece mediante webhooks HTTP. Cuando el usuario carga una imagen y solicita el análisis, los datos se encapsulan y transmiten al nodo receptor del flujo de trabajo, iniciando una secuencia de procesamiento estructurada en las siguientes fases:

- **Ingesta y normalización:** El flujo inicia con la recepción de la información, seguida de una etapa de configuración y normalización de datos para asegurar la compatibilidad con los modelos subsecuentes.
- **Orquestación y validación sísmica:** La información ingresa a un nodo de decisión denominado “*Orquestador Sísmico*”. Este agente utiliza un modelo de OpenAI para inspeccionar la naturaleza del archivo de entrada. El modelo evalúa si el archivo cumple con el formato requerido y verifica semánticamente si la imagen corresponde a una línea sísmica válida. En caso de no superar esta validación, el proceso se interrumpe.
- **Preprocesamiento de imagen:** Las imágenes validadas se someten a una conversión a escala de grises, seguida de la aplicación del algoritmo CLAHE (*Contrast Limited Adaptive Histogram Equalization*). Este paso es crítico para maximizar la nitidez y resaltar los reflectores sísmicos antes de la segmentación.
- **Segmentación de sismofacies:** La imagen mejorada alimenta al modelo de segmentación semántica DeepLabv3. Este modelo procesa la entrada y genera como salida una máscara de segmentación, donde cada sismofacie identificada se representa mediante una codificación de colores específica.
- **Interpretación geológica automatizada:** Tras una verificación automática de la calidad de la segmentación, el sistema activa un nodo LLM (OpenAI)⁹. Utilizando un *prompt* técnico predefinido, el modelo analiza las características detectadas y genera una interpretación geológica textual detallada de las sismofacies.
- **Generación de reporte técnico:** Dado que el LLM genera texto y el modelo de segmentación genera imágenes, se emplea un nodo de fusión (*merge node*) para consolidar ambos resultados. Finalmente, el sistema estructura esta información híbrida para la generación automática de un reporte técnico en formato PDF, el cual es devuelto a la interfaz de Streamlit para su descarga inmediata por parte del usuario.
- **Integración de algoritmos:** La ejecución de los algoritmos de visión por computadora y segmentación se externalizó mediante una arquitectura de servicios en la nube. El código fuente, que abarca desde la conversión a escala de grises y la mejora de nitidez (CLAHE) hasta la implementación de la arquitectura DeepLabv3, fue desarrollado en un entorno de Google Colab PRO¹⁰ utilizando bibliotecas de Python¹¹. No obstante, debido a la naturaleza efímera de las sesiones en Colab, que impide una conexión constante, se implementó un despliegue mediante Modal¹². Esta plataforma encapsula el entorno de ejecución y lo expone a través de un endpoint HTTP persistente (URL estática), permitiendo que el

⁷<https://n8n.io>

⁸<https://streamlit.io>

⁹<https://openai.com>

¹⁰<https://colab.research.google.com>

¹¹<https://www.python.org>

¹²<https://modal.com>

orquestador n8n realice peticiones de inferencia estables al modelo remoto y supere las limitaciones de conectividad de los entornos interactivos. De esta forma, se cierra el ciclo de procesamiento devolviendo a la aplicación web los resultados del análisis en un formato estructurado (JSON), el cual integra las rutas de las imágenes generadas junto con el texto interpretativo para su visualización final.

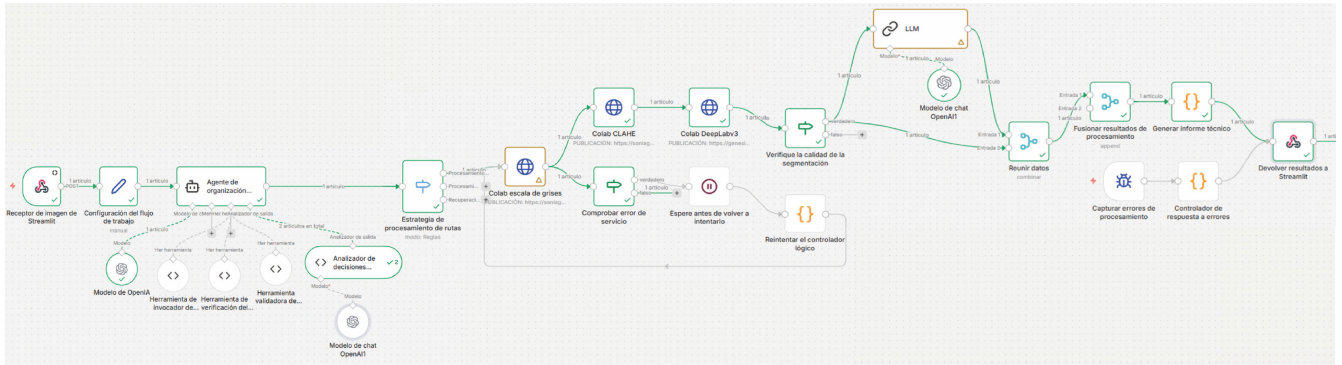


Figura 10: Workflow en n8n.

3. Experimentación

Esta sección presenta el desarrollo experimental del sistema propuesto y los resultados obtenidos a partir del entrenamiento y evaluación del modelo de segmentación automática de sismofacies descrito en la sección anterior. En primer lugar, se detallan las características de la plataforma computacional empleada y la preparación del conjunto de datos. Posteriormente, se describen las condiciones de entrenamiento, el comportamiento del modelo a través de las curvas de aprendizaje y la evaluación final sobre el conjunto de prueba.

3.1. Plataforma computacional

El entrenamiento y la evaluación del modelo se realizaron utilizando la plataforma Google Colab PRO, la cual fue seleccionada debido a su accesibilidad, facilidad de uso y disponibilidad de recursos computacionales orientados a aplicaciones de aprendizaje profundo. Esta plataforma permite ejecutar código en un entorno basado en la Web, sin requerir configuraciones locales complejas.

El entorno computacional empleado cuenta con un sistema operativo Linux Ubuntu, un procesador Intel, Xeon, una unidad de procesamiento gráfico (GPU) NVIDIA Tesla con 16 GB de memoria dedicada y una memoria RAM del sistema superior a 50 GB. En google colab se llevaron a cabo en un entorno de computación acelerada por GPU, utilizando Python como lenguaje de programación principal y el framework PyTorch para el desarrollo y entrenamiento de redes neuronales profundas. El modelo fue ejecutado con soporte de CUDA 11.8, permitiendo el uso de procesamiento paralelo en GPU. El sistema empleó una GPU NVIDIA A100 compatible con CUDA, utilizada para acelerar tanto el entrenamiento como la inferencia del modelo DeepLabv3. Estas características permitieron entrenar modelos de redes neuronales convolucionales profundas de manera eficiente.

La implementación se realizó en el lenguaje de programación Python, utilizando las librerías TensorFlow y Keras para la construcción y entrenamiento del modelo DeepLabv3, así como bibliotecas complementarias como NumPy para el manejo de arreglos numéricos, OpenCV para el procesamiento de imágenes sísmicas y Matplotlib para la visualización de resultados.

3.2. Entrenamiento del modelo de segmentación automática de sismofacies

El entrenamiento del modelo se llevó a cabo bajo un enfoque de aprendizaje automático, en el cual el objetivo es aprender la correspondencia entre las imágenes sísmicas de entrada y sus respectivas máscaras de segmentación. Durante este proceso, el modelo ajusta de manera iterativa sus parámetros internos minimizando la función de pérdida definida.

Para mejorar la precisión en la segmentación de sismofacies, se implementó la técnica de transfer learning. Esta estrategia permite reutilizar el conocimiento previamente adquirido por un modelo entrenado en grandes volúmenes de datos y adaptarlo a una tarea específica, evitando así la necesidad de entrenarlo desde cero y reduciendo considerablemente los requerimientos computacionales y de datos. El modelo seleccionado fue DeepLabv3, previamente entrenado con miles de imágenes genéricas para tareas de segmentación semántica.

Hablar de las 14 clases consideradas

En esta etapa, el modelo fue afinado (transfer learning) utilizando el conjunto de imágenes sísmicas previamente procesadas y etiquetadas. El entrenamiento se configuró bajo un esquema de segmentación semántica multiclase, considerando un total de 14 clases de sismofacies previamente definidas en la sección correspondiente. En consecuencia, la capa de salida del modelo fue adaptada para generar 14 mapas de probabilidad, uno por cada categoría, permitiendo la clasificación pixel a pixel de las distintas facies presentes en las líneas sísmicas. De esta manera, los pesos de la red se ajustaron para que el modelo aprendiera las características particulares asociadas a cada clase, mejorando su capacidad de discriminación y precisión en la segmentación. Dentro de cada carpeta de imágenes y máscaras se crean las subcarpetas de train, val y test, las imágenes y etiquetas de la carpeta train se usan para entrenar y afinar el modelo, en cada iteración del entrenamiento el modelo es validado con las imágenes de la carpeta Val, y, al final, con las imágenes de la carpeta test, se realiza la prueba del modelo con imágenes no vistas.

Antes del entrenamiento se configuró los siguientes hiperparámetros: un máximo de 50 épocas, utilizando el optimizador Adam y la función de pérdida *Sparse Categorical Crossentropy* y una tasa de aprendizaje de 0,001. Además, se implementó el mecanismo de *Early Stopping*, el cual permite detener el entrenamiento cuando la métrica de validación deja de mejorar durante varias épocas consecutivas, evitando el sobreajuste del modelo.

Una vez configurados los hiperparámetros, se inicia el proceso de entrenamiento del modelo. El modelo DeepLabv3 es cargado con pesos preentrenados y alimentado con las imágenes correspondientes al conjunto de entrenamiento, las cuales son procesadas a través de la red para generar predicciones asociadas a cada clase definida. A lo largo del entrenamiento, se evalúan de manera continua los errores producidos por la red tanto en el conjunto de entrenamiento como en el de validación, a partir de los cuales se calculan los valores de pérdida mediante el proceso de backpropagation. Este procedimiento se ejecuta de forma iterativa durante 50 épocas, con el objetivo de minimizar el error y optimizar el desempeño del modelo.

3.3. Análisis de las curvas de aprendizaje

Tras completar el proceso de entrenamiento, el modelo DeepLabv3 adquiere la capacidad de identificar y diferenciar automáticamente las distintas sismofacies presentes en las imágenes sísmicas. El desempeño del aprendizaje se evalúa mediante el análisis de las curvas de entrenamiento, las cuales representan la evolución de la función de pérdida (loss) y de la precisión (accuracy) tanto en el conjunto de entrenamiento como en el conjunto de validación.

La Figura 12 presenta dichas curvas, donde se observa un incremento sostenido de la precisión a lo largo de las épocas, lo que evidencia un comportamiento estable y adecuado del modelo durante el entrenamiento. Asimismo, la diferencia moderada entre las curvas de entrenamiento y validación sugiere un equilibrio apropiado entre la capacidad de ajuste y la generalización. La ausencia de separaciones marcadas o comportamientos irregulares entre ambas curvas indica que el modelo no presenta problemas de sobreajuste (overfitting) ni de subajuste (underfitting),

confirmando que el proceso de entrenamiento se desarrolló de manera correcta y consistente. Las métricas obtenidas de Accuracy de entrenamiento final: 0.891: y Accuracy de validación final: 0.863

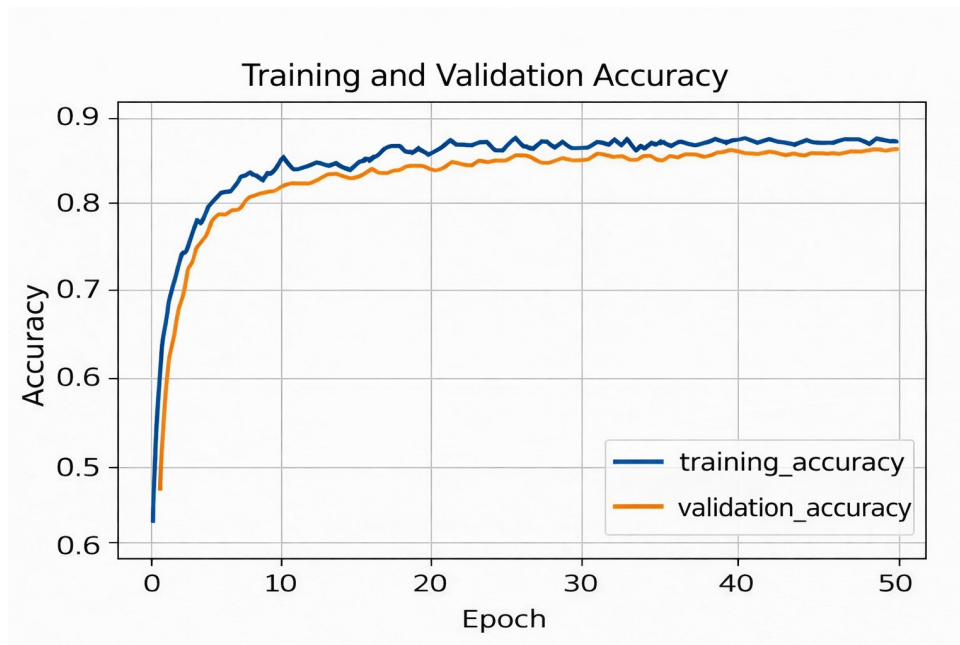


Figura 11: Curvas de aprendizaje del modelo DeepLabv3: (a) pérdida y (b) precisión para los conjuntos de entrenamiento y validación.

3.4. Evaluación del modelo

Una vez concluido el proceso de entrenamiento, el modelo fue evaluado empleando el conjunto de datos de prueba con imágenes sísmicas que no formaron parte del entrenamiento ni de la validación, con el fin de analizar de manera objetiva su capacidad de generalización y su desempeño en condiciones cercanas a un escenario real de aplicación. Para esta etapa se utilizó el modelo que alcanzó el mejor rendimiento durante el entrenamiento, el cual fue fortalecido mediante la aplicación de técnicas de datos. Éstas precisiones y sensibilidad (Recall) indican una adecuada capacidad para identificar correctamente las regiones pertenecientes a cada facies y reducir errores de clasificación.

El análisis detallado por clase (Figura 13) revela que el modelo presenta un desempeño particularmente alto en las clases 7 a 14, alcanzando máximos cercanos a 0.90. Este comportamiento sugiere que dichas sismofacies poseen patrones sísmicos más definidos y diferenciables, lo que facilita su segmentación automática. En contraste, las clases con valores moderados, especialmente en los primeros niveles, reflejan la complejidad inherente a la segmentación de facies con características texturales y estructurales similares, así como posibles diferencias en la distribución y representatividad de estas clases dentro del conjunto de datos. No obstante, la cercanía entre los valores de precisión y recall en la mayoría de las clases indica que el modelo mantiene un equilibrio adecuado entre falsas detecciones y omisiones, evitando sesgos significativos hacia la sobresegmentación o subsegmentación. En conjunto, los resultados obtenidos en el conjunto de prueba son coherentes con los observados durante la fase de validación, lo que confirma que el modelo DeepLab V3 presenta una buena capacidad de generalización y un desempeño sólido para la segmentación de sismofacies a partir de imágenes sísmicas, validando su potencial aplicación en estudios geológicos y geofísicos reales.

Clase	Precision	Recall
1	0.6358	0.6912
2	0.6911	0.7119
3	0.6805	0.6749
4	0.7181	0.7361
5	0.7408	0.7008
6	0.7532	0.7424
7	0.8273	0.8073
8	0.8999	0.8386
9	0.8631	0.8512
10	0.8918	0.8978
11	0.9108	0.8908
12	0.9346	0.9046
13	0.9466	0.9276
14	0.9106	0.9172

Figura 12: Métricas del modelo sobre el conjunto de datos de test

El modelo demostró un desempeño consistente en la detección e interpretación automática de sismofacies, generando máscaras coherentes espacialmente y acordes con las interpretaciones manuales. La Figura 13 muestra un ejemplo representativo de los resultados obtenidos en comparación con una interpretación tradicional.

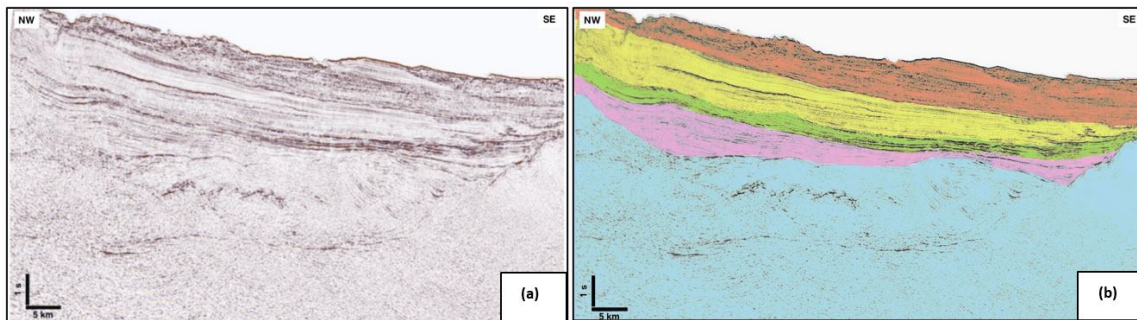


Figura 13: Ejemplo de resultados de segmentación automática: (a) imagen sísmica original y (b) máscara de segmentación generada por el modelo.

4. Resultados

El producto final está diseñado para cumplir con requerimientos fundamentales como facilidad de uso, accesibilidad e implementación de una interfaz intuitiva que permita al usuario automatizar la interpretación de líneas sísmicas. En función de estos criterios, se optó por desplegar el modelo entrenado mediante una aplicación web accesible desde cualquier navegador.

La plataforma GeoSeismicAI (Figura 14) presenta en su página principal la carga de la imagen sísmica y posteriormente, la clasificación automática de sismofacies junto con una interpretación geológica preliminar generada por el modelo. Además, permite exportar tanto la imagen segmentada como la información asociada en un documento estructurado con formato de reporte técnico.

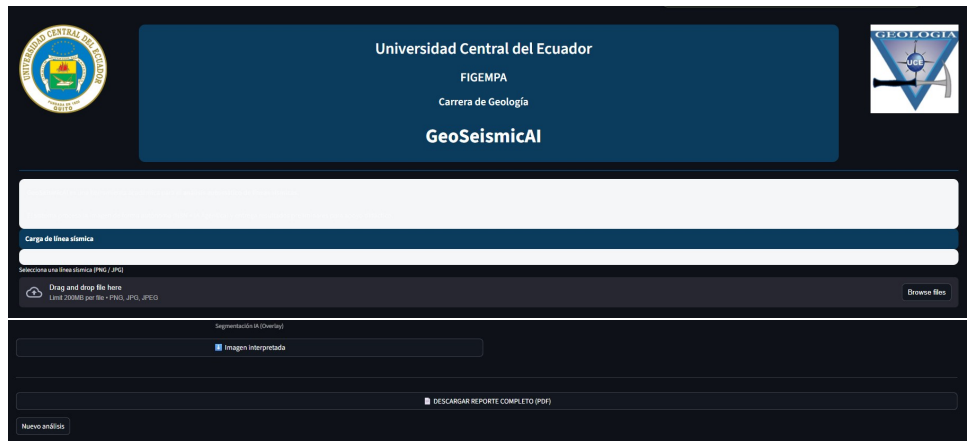


Figura 14: Página principal de GeoSeismicAI.

4.1. Aplicación web

El despliegue de la plataforma GeoSeismicAI se realizó mediante la integración de GitHub como sistema de control de versiones y Streamlit como framework para el desarrollo y publicación de la aplicación web de manera interactiva. Esta arquitectura permitió garantizar reproducibilidad, accesibilidad y mantenimiento continuo del sistema desarrollado.

La plataforma integra la aplicación desarrollada en Streamlit a través de un enlace directo, lo que permite al usuario cargar imágenes sísmicas y obtener una interpretación en tiempo real. Adicionalmente, el sistema ofrece opciones de descarga tanto de la imagen interpretada en formato .jpg como de un reporte final en formato .pdf, el cual incluye una comparativa entre la imagen original, las sismofacies segmentadas y una interpretación geológica preliminar.

En la Figura 15 se muestra una representación general del sitio Web y una captura de pantalla de la aplicación en funcionamiento.

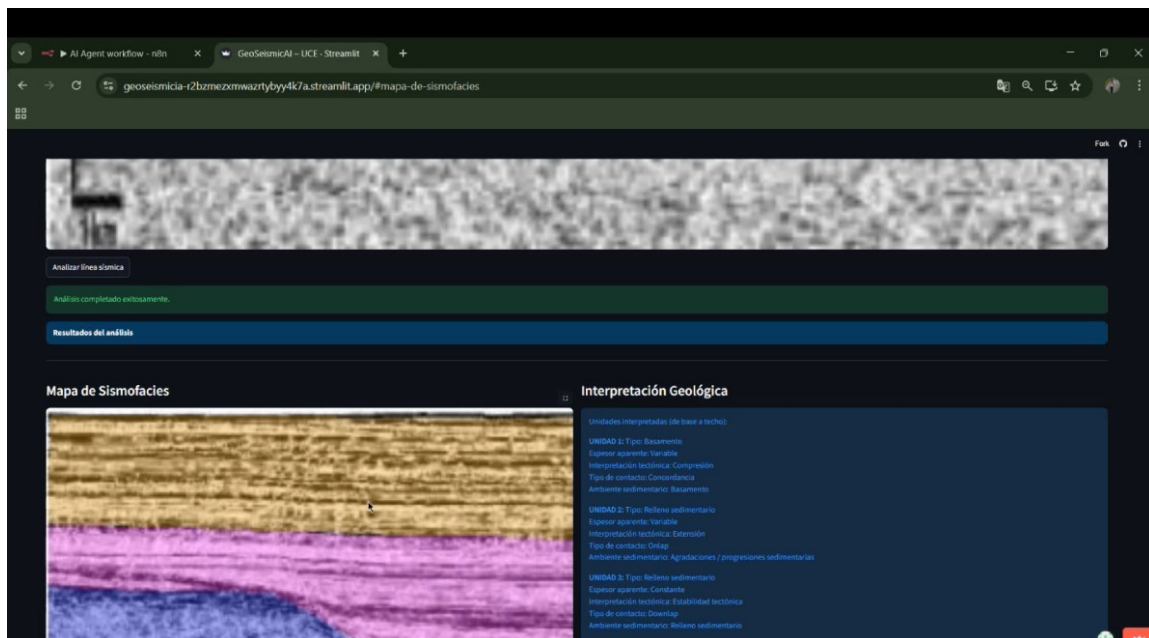
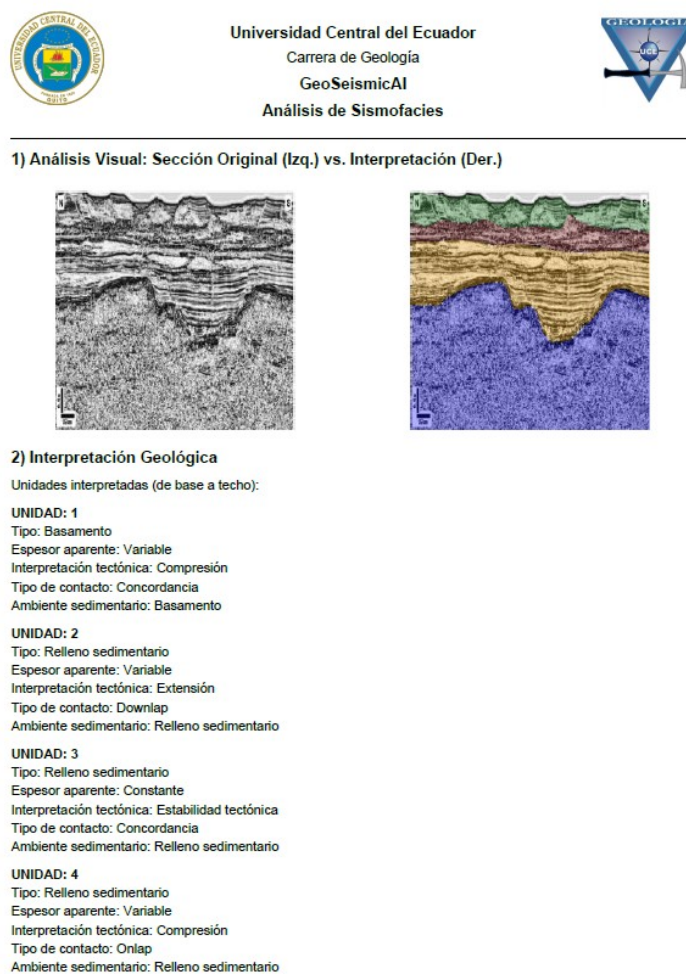


Figura 15: Vista general del sitio Web y ejemplo de uso de la aplicación para la segmentación automática de sismofacies a partir de una imagen sísmica.

4.2. Reporte técnico

El sistema GeoSeismicAI consolida los resultados del procesamiento en un reporte técnico estandarizado en formato PDF, diseñado para integrar la evidencia visual con la interpretación geológica. Este documento presenta una comparativa horizontal de la línea sísmica original y la imagen interpretada, aplicando sobre esta última una técnica de superposición (overlay) con transparencia alfa ($\alpha = 0,3$) que permite visualizar la máscara de sismofacies generada por la IA sin ocultar la textura de los reflectores subyacentes. De manera complementaria, el reporte incorpora la interpretación geológica textual producida por el modelo de lenguaje (LLM), ofreciendo una síntesis narrativa que identifica unidades estratigráficas, geometrías de los reflectores y ambientes deposicionales probables.



Continúa en la siguiente página...

Figura 16: Ejemplo de reporte técnico generado por GeoSeismicAI.

5. Conclusiones

- Se desarrolló un sistema para la detección y segmentación automática de sismofacies en imágenes sísmicas bidimensionales mediante técnicas de inteligencia artificial y aprendizaje profundo. El dataset se construyó inicialmente a partir de 200 imágenes sísmicas etiquetadas manualmente y fue ampliado hasta un total de 340 imágenes mediante técnicas de aumento de datos.

- La aplicación de procesos de preprocesamiento, como la conversión a escala de grises y la mejora de contraste mediante CLAHE, resultó fundamental para realzar los reflectores sísmicos y optimizar el desempeño del modelo, alcanzando una precisión aproximada del 89 %.
- El modelo DeepLabv3 demostró ser una arquitectura adecuada para la segmentación de sismofacies, debido a su capacidad para capturar información multiescala y preservar el contexto espacial de las imágenes sísmicas, logrando resultados coherentes y reproducibles.
- La interpretación generada por el sistema constituye una aproximación preliminar válida que contribuye a mejorar la productividad del geólogo en las etapas tempranas de exploración, sin reemplazar el criterio del especialista, sino funcionando como una herramienta de apoyo en el análisis sísmico.

6. Recomendaciones

- Ampliar el conjunto de imágenes sísmicas utilizadas en el entrenamiento del modelo, incorporando diferentes contextos geológicos y condiciones de adquisición, con el fin de mejorar la capacidad del sistema para reconocer patrones visuales asociados a anomalías de hidrocarburos.
- Profundizar en la aplicación de técnicas de detección de objetos sobre imágenes sísmicas, permitiendo una localización más precisa de las zonas anómalas y una mejor delimitación de las áreas de interés.
- Automatizar la identificación de estructuras geológicas, como fallas o discontinuidades, que continúen aportando al proceso de interpretación geológica en trabajos futuros.

7. Datos disponibles y código

Las imágenes sísmicas utilizadas en este estudio corresponden a secciones sísmicas en formato de imagen, preparadas y preprocesadas para su uso en tareas de visión por computador. Debido a restricciones de confidencialidad y derechos de uso, el conjunto de datos no se encuentra disponible de forma pública; sin embargo, puede ser solicitado previa petición por correo electrónico.

Agradecimientos

Deseamos expresar nuestro más sincero agradecimiento al Ing. Cristian Zura por su guía a lo largo del desarrollo de este trabajo. Su apoyo fue importante no solo por la provisión de las imágenes sísmicas utilizadas en el estudio, sino también por sus explicaciones, orientación técnica y disposición permanente para aclarar conceptos relacionados con la interpretación sísmica y el análisis del subsuelo.

De igual manera, se agradece a Dalember Vallejo por su acompañamiento, consejos oportunos y apoyo constante, especialmente en momentos críticos del proyecto en los que las decisiones técnicas y metodológicas resultaban complejas. Su ayuda y motivación nos ayudó a superar dificultades y culminar satisfactoriamente este trabajo.

Referencias

- [1] Sheriff, R. E., & Geldart, L. P. (1995). *Exploration Seismology*. Cambridge University Press.

- [2] Yilmaz, Ö. (2001). *Seismic Data Analysis: Processing, Inversion, and Interpretation of Seismic Data*. SEG.
- [3] Veeken, P. C. (2007). *Seismic Stratigraphy, Basin Analysis and Reservoir Characterisation*. Elsevier.
- [4] Misra, A. A., & Mukherjee, S. (Eds.). (2018). *Atlas of Structural Geological Interpretation from Seismic Images*. Wiley-Blackwell.
- [5] Kearey, P., Brooks, M., & Hill, I. (2002). *An Introduction to Geophysical Exploration*. Blackwell Science.
- [6] Steeples, D. W. (2001). Engineering and environmental geophysics at the millennium. *Geophysics*, 66(1), 31–35.
- [7] Brown, A. R. (2011). *Interpretation of Three-Dimensional Seismic Data*. SEG Books.
- [8] Bond, C. E. (2015). Uncertainty in structural interpretation: Lessons to be learnt. *Journal of Structural Geology*, 74, 85–102.
- [9] Chopra, S., & Marfurt, K. J. (2007). *Seismic Attributes for Prospect Identification and Reservoir Characterization*. SEG.
- [10] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- [11] Araya-Polo, M., Dahlke, T., Frogner, C., Zhang, C., Poggio, T., & Hohl, D. (2017). Automated fault detection without seismic processing. *The Leading Edge*, 36(3), 208–214.
- [12] Huang, L., Dong, X., & Clee, T. E. (2017). A scalable deep learning platform for identifying geologic features. *The Leading Edge*, 36(3), 249–256.
- [13] Zeng, Y., Jiang, K., & Chen, J. (2018). Automatic seismic salt interpretation with deep convolutional neural networks. *arXiv preprint arXiv:1812.01101*.
- [14] Babakhin, Y., Sanakoyeu, A., & Kitamura, H. (2019). Semi-supervised segmentation of salt bodies in seismic images using an ensemble of convolutional neural networks. *arXiv preprint arXiv:1904.04445*.
- [15] Zhang, H., Zhu, P., & Liao, Z. (2023). SaltISNet3D: Interactive salt segmentation from 3D seismic images using deep learning. *Remote Sensing*, 15(9), 2319.
- [16] Alfarraj, M., AlRegib, G., & Deriche, M. (2024). KG-Unet: A knowledge-guided deep learning approach for seismic facies segmentation. *Earth Science Informatics*, 17, 1967–1981.
- [17] Di, H., Wang, X., & Chen, Y. (2024). Literature review on deep learning for the segmentation of seismic images. *Earth-Science Reviews*, 258, 104955.
- [18] Lima, G., Ramos, G., Rigo, S., Zeiser, F., & da Silveira, A. (2020). Binary segmentation of seismic facies using encoder-decoder neural networks. *arXiv preprint arXiv:2012.03675*.
- [19] Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*.
- [20] Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In P. S. Heckbert (Ed.), *Graphics Gems IV* (pp. 474–485). Academic Press.
- [21] Aribido, O. J., AlRegib, G., & Alaudah, Y. (2021). Self-supervised delineation of geological structures using orthogonal latent space projection. *arXiv preprint arXiv:2108.09605*.

- [22] Alfarraj, M., AlRegib, G., & Deriche, M. (2024). Literature review on deep learning techniques for salt dome segmentation in seismic images. *Journal of Applied Geophysics*, 230, 105504.
- [23] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention Is All You Need*. In *Advances in Neural Information Processing Systems* (Vol. 30, pp. 5998–6008). Long Beach, CA, USA.
- [24] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. arXiv preprint arXiv:2010.11929. Disponible en <https://arxiv.org/abs/2010.11929>
- [25] Kerner, H. R., Bell III, J. F., & Amor, H. B. (2018). Context-dependent image quality assessment of Mars images using convolutional neural networks. *Computers & Geosciences*, 118, 109–121.
- [26] Palafox, L. F., Hamilton, C. W., Scheidt, S. P., & Alvarez, A. M. (2017). Automated detection of geological landforms on Mars using convolutional neural networks. *Computers & Geosciences*, 101, 48–56.
- [27] Shi, Y., Wu, X., & Fomel, S. (2018). Automatic salt-body classification using deep-convolutional neural network. In *SEG Technical Program Expanded Abstracts 2018* (pp. 1971–1975).
- [28] Waldeland, A. U., Jensen, A. C., Gelius, L.-J., & Solberg, A. H. S. (2018). Convolutional neural networks for automated seismic interpretation. *The Leading Edge*, 37(7), 529–537.
- [29] Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., & Wen, J.-R. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), 186345.
- [30] Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., & Zhang, Y. (2023). The rise and potential of large language model based agents: A survey. *arXiv preprint arXiv:2309.07864*.
- [31] Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. M. (2017). Enhanced Deep Residual Networks for Single Image Super-Resolution. *CVPR Workshops*.
- [32] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of CVPR*, 3431–3440.

Sobre los autores

Silvana Marjorie Hermosa



Estudiante de Geología en la Universidad Central del Ecuador. Ha realizado pasantías en EP PETROECUADOR, en el activo Lago Agrio-Libertador, así como en PetroTech, en el Instituto de Investigación Geológico y Energético (IIGE), en el ECU 911, desempeñándose en áreas de modelamiento geológico, cartografía regional, estadística y análisis de datos. Cuenta con formación en registros eléctricos, sedimentología, petrografía, sistemas de información geográfica y estadística, y ha participado en cursos especializados de interpretación de registros de pozos, incluyendo el curso “Preparación inteligente de registros de pozos para estudios integrados de yacimientos (Norte y Sur América)”, donde se aplican técnicas de inteligencia artificial para el análisis e integración de datos geológicos y geofísicos. Sus intereses académicos se centran en la aplicación de inteligencia artificial y aprendizaje profundo en la interpretación sísmica, registros de pozos y modelamiento geológico.

Génesis Campos Duicela



Estudiante de 10mo semestre de Ingeniería en Geología en la Universidad Central del Ecuador, con formación orientada al análisis geológico aplicado. Realizó pasantías preprofesionales en la Agencia de Regulación y Control Minero (ARCOM), donde adquirió experiencia en revisión técnica de información geológica, normativa minera y gestión de datos institucionales. Posteriormente, desarrolló pasantías en el Centro de Investigación Quito - EP PETROECUADOR, en el área de investigación, fortaleciendo conocimientos en análisis de información del sector hidrocarburífero, apoyo a procesos petrográficos y generación de informes técnicos. Asimismo, ha participado en actividades académicas y técnicas vinculadas al Instituto de Investigación Geológico y Energético (IIGE), complementando su formación en geología aplicada, además, cuenta con una participación activa en asociaciones técnicas como la AAPG – Capítulo Estudiantil UCE y la Asociación de Ingenieros de Minas del Ecuador (AIME). Su trayectoria refleja un interés constante por la investigación aplicada y el trabajo interdisciplinario.

Christian Mejia-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.

Sonia Gualán Cartuche



Estudiante de 10mo semestre de Ingeniería en Geología en la Universidad Central del Ecuador. Ha desarrollado experiencia académica y práctica en interpretación geológica, análisis petrográfico, petrofísica y elaboración de informes técnicos. Realizó prácticas preprofesionales en el Instituto de Investigación Geológico y Energético (IIGE), participando en cartografía geológica y planificación mediante SIG, y en EP Petroecuador, en el CIQ (Centro de Investigaciones Quito) colaborando en ensayos de porosidad y permeabilidad, ensayos de daños a la formación y elaboración de informes. Además, cuenta con interés y experiencia en el uso de Inteligencia Artificial aplicada a la geología, especialmente en interpretación automatizada de datos sísmicos mediante técnicas de aprendizaje profundo. Complementa su formación con participación en proyectos ambientales y sociales enfocados en sostenibilidad.

Ronny López Tasiguano



Estudiante de 10mo semestre de la carrera de Ingeniería en Geología en la Universidad Central del Ecuador, con formación orientada a la geología aplicada. Geólogo de campo participando en levantamiento geológico para el proyecto vial Warintz en Morona Santiago. Realizó prácticas preprofesionales en la empresa Terrahidro, donde adquirió experiencia en la ejecución de ensayos de penetración estándar (SPT), apertura de calicatas, muestreo y descripción geotécnica de suelos, así como en actividades relacionadas con la construcción y evaluación de pozos hidrogeológicos. Posteriormente, desarrolló prácticas en el Centro de Investigaciones Quito, EP PETROECUADOR, en el área de investigación, fortaleciendo conocimientos en petrografía, análisis mediante coregamma espectral y microscopía electrónica de barrido (MEB), además de brindar soporte técnico en la elaboración de informes para concesiones mineras y gestión de riesgos geológicos. Actualmente, participa activamente en el concurso Imperial Barrel Award.