

---

# CNN + LLM: INTERPRETACIÓN BÁSICA DEL SUBSUELO USANDO IMÁGENES HVSR

---

A PREPRINT

Katherine Ayala

Renata Bohórquez

Eduardo Guerrero

Víctor Solís

Christian Mejia-Escobar

Facultad de Ingeniería en Geología, Minas, Petróleos y Ambiental (FIGEMPA)  
Universidad Central del Ecuador  
Quito, Ecuador  
{kvayala,rpbohorquez,eggerrero,vhsolis,cimejia}@uce.edu.ec

20 de octubre de 2025

## ABSTRACT

Las gráficas HVSR (Relación Espectral Horizontal-Vertical) proporcionan información valiosa para una caracterización preliminar del subsuelo, útil en estudios geotécnicos, microzonificación y evaluación sísmica, planificación urbana, entre otras áreas. El análisis y la interpretación de estas gráficas demandan tiempo y dependen altamente del experto, lo que puede provocar errores y subjetividad. El presente estudio propone automatizar este proceso mediante inteligencia artificial basada en técnicas de visión y lenguaje. A partir de la selección de 90 imágenes HVSR que fueron preprocesadas para extraer la curva espectral dominante, se formó un conjunto de entrenamiento para una red neuronal convolucional que predice los parámetros espectrales de frecuencia fundamental ( $F_0$ ) y amplitud máxima ( $A_0$ ). Estas predicciones permiten el cálculo del índice de vulnerabilidad sísmica  $K_g$ , lo que se integra en un prompt que es procesado por un modelo de lenguaje (LLM) para generar descripciones básicas y rápidas del subsuelo. Las predicciones mostraron una alta correlación con los valores interpretados por un experto, alcanzando valores bajos de Error Absoluto Medio (MAE) y Error Cuadrático Medio (MSE). Proporcionamos un sistema multimodal de texto, imagen y voz para la interpretación automatizada de gráficas HVSR que constituye una herramienta confiable y accesible para estudiantes, profesionales y usuarios no especialistas, mejorando significativamente la velocidad y objetividad del análisis geofísico.

**Keywords** HVSR, CNN, LLM, interpretación, subsuelo, geofísica

## 1. Introducción

¿Se puede enseñar a una computadora a describir la Tierra? Suena a ciencia ficción, pero ahora es una realidad. La interpretación del subsuelo mediante métodos geofísicos es fundamental en la toma de decisiones en diversos ámbitos, como la planificación territorial, la evaluación de riesgos naturales, la exploración de recursos naturales y el diseño de infraestructuras [1][2]. Una técnica popular para este propósito es el HVSR (Horizontal to Vertical Spectral Ratio), que compara los espectros horizontales y verticales del movimiento del suelo registrados durante microtemores (vibración ambiental) para determinar la frecuencia de resonancia del sitio [3]. En términos simples, se escucha el latido del suelo para conocer su rigidez y espesor [4]. Para tal fin, se utilizan geófonos orientados en las tres direcciones del espacio: dos horizontales (N-S y E-O) y una vertical (Z) [5].

En comparación con otros métodos geofísicos, HVSR destaca por su simplicidad, rapidez, eficacia, bajo costo y carácter no invasivo. Diversos estudios validan la aplicación de esta técnica en múltiples campos, entre los que se encuentran: la evaluación del terreno, estimación de espesores y condiciones de licuefacción del suelo en geotecnia y minería [6]; la detección de estructuras geológicas ocultas de tipo tectónico [7] y estructuras ciegas, como fallas inversas u oblicuas

[8]; la caracterización preliminar de zonas inestables en áreas de difícil acceso [2]; la identificación de superficies de ruptura y evaluación del espesor de los sedimentos movilizados en un deslizamiento [9]; el análisis de vulnerabilidad, clasificación y microzonificación sísmica [3]; la estimación de parámetros geotécnicos [10]; la gestión de riesgos geológicos [11] y el ordenamiento territorial tanto urbano como rural [12].

Los trabajos mencionados evidencian la utilidad de HVSR como una técnica preliminar eficaz para la exploración y caracterización del suelo y subsuelo. Sin embargo, uno de los grandes retos de esta técnica es la interpretación de los resultados. Esta interpretación se centra en la gráfica del HVSR, donde se debe hallar un pico muy marcado que corresponde a los valores máximos del espectro de amplitud ( $A_0$ ) y la frecuencia natural de resonancia del sitio ( $f_0$ ) [13]. En términos generales, el análisis persigue identificar las curvas y los picos principales en los espectros.

La interpretación precisa de estas imágenes depende de humanos expertos que evalúan múltiples factores y parámetros. Aunque este método es eficaz, su aplicación resulta altamente laboriosa, requiere gran cantidad de tiempo, incorpora subjetividad y, en ocasiones, ambigüedad o inconsistencias. Estos errores pueden conducir a interpretaciones incorrectas, desencadenando en una toma de decisiones equivocada [14]. Además, limita la capacidad de procesamiento de grandes volúmenes de datos. En otras palabras, si no se cuenta con experiencia y tiempo suficientes que permitan analizar una gran cantidad de imágenes, la tarea se extenderá considerablemente.

En vista de estos desafíos, este estudio propone el uso de inteligencia artificial (IA), la cual ofrece alternativas potencialmente más eficientes a los métodos tradicionales [15]. En los últimos años, varios estudios documentan aportes relevantes de la IA para el tratamiento de problemas geológicos y geotécnicos. Específicamente, en [16] se presenta el algoritmo AutoHVSR, que automatiza completamente el análisis de datos HVSR provenientes de microtemores o registros sísmicos, empleando parámetros definidos por el usuario y modelos de aprendizaje automático entrenados con 1108 mediciones. Sin embargo, este algoritmo no realiza una interpretación del terreno automatizada.

Una interpretación automática de las gráficas HVSR debe integrar dos funcionalidades esenciales: el reconocimiento de características en imágenes, y su posterior descripción textual y contextual. Existen estudios que han demostrado el potencial de los modelos de deep learning de visión y lenguaje para estas tareas en el ámbito geológico. Por ejemplo, la aplicación del aprendizaje no supervisado en la clasificación de facies en imágenes de rocas carbonatadas [17], las Redes Neuronales Artificiales (ANN) en la predicción de la resistencia al corte del suelo [18], las Redes Neuronales Convolucionales (CNN) con Sliding Windows, Pirámide Gaussiana y Non-Maxima Suppression en la detección de características para clasificación [19], el uso de LLaVA (CLIP + Vicuna) en la comprensión visual-lingüística multimodal (ScienceQA) [20].

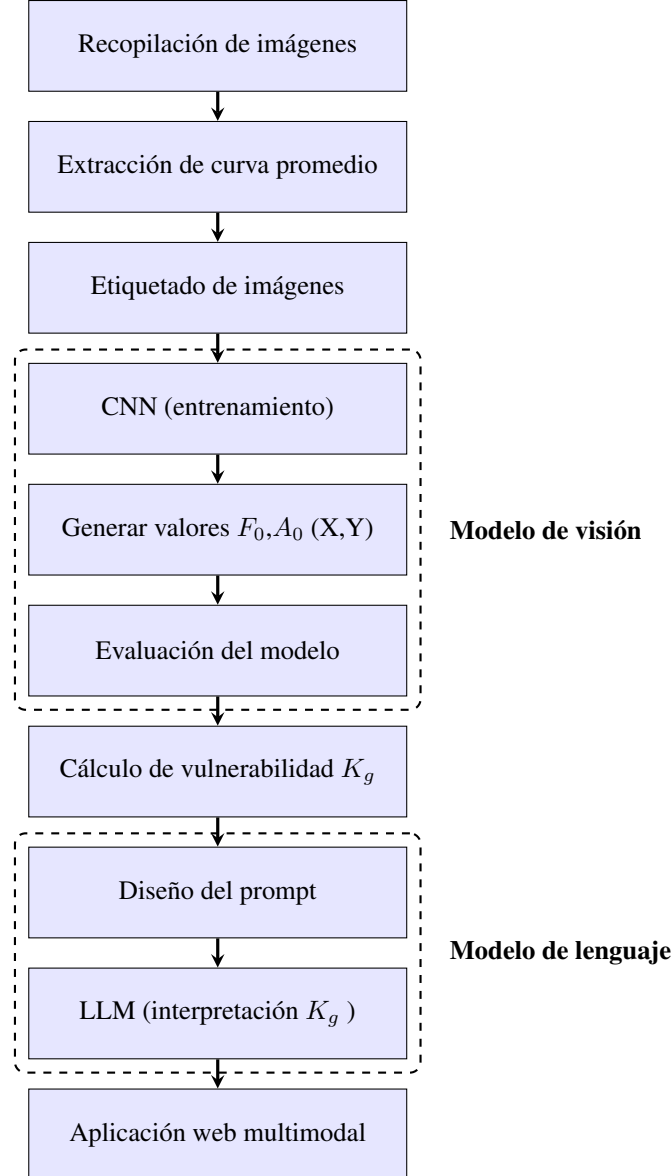
Combinar técnicas de visión por computadora y procesamiento de lenguaje natural forma parte de la tendencia emergente en IA multimodal, donde distintos tipos de datos como imágenes y texto son tratados de manera coordinada para mejorar la comprensión y toma de decisiones [21]. Aunque un Modelo de Lenguaje Grande Multimodal (MLLM) es capaz de interpretar y describir imágenes de propósito general, las gráficas HVSR constituyen un tipo de imagen altamente específica del dominio geofísico. Un tratamiento preciso de estas imágenes especializadas demanda reentrenar o ajustar el MLLM, lo cual es extremadamente costoso y lento.

Por esta razón, la estrategia propuesta implementa una CNN para estimar parámetros fundamentales del terreno a partir de gráficas HVSR. Estos valores permiten calcular un índice de vulnerabilidad que se integra en un prompt cuidadosamente diseñado, el cual es procesado por un modelo de lenguaje para generar una interpretación geofísica preliminar del subsuelo. El buen desempeño de esta solución ha motivado el desarrollo de una aplicación disponible en línea que genera de manera textual y verbal la interpretación de gráficas HVSR. Esta herramienta constituye un aporte significativo en la interpretación de los espectros de frecuencia, obteniendo información valiosa sobre el terreno, con el objetivo de optimizar el análisis y la toma de decisiones en la geotecnia, la geología y la gestión de riesgos.

El contenido del documento incluye una visión general del proyecto y trabajos relacionados en la Sección 1; la Sección 2 describe la metodología aplicada; la Sección 3 explica la parte experimental, así como los resultados obtenidos; la Sección 4 se dedica al desarrollo de la aplicación multimodal. Para finalizar, las conclusiones son expuestas en la Sección 5.

## 2. Metodología

Este proyecto tiene como objetivo desarrollar un sistema automatizado de reconocimiento e interpretación de imágenes de curvas HVSR mediante técnicas de deep learning, con el fin de optimizar el análisis sísmico pasivo en estudios de caracterización del subsuelo. La principal contribución radica en automatizar un proceso tradicionalmente manual, reduciendo la subjetividad y el tiempo requerido, además de validar la calidad de las mediciones. Esta herramienta inteligente permitirá aplicar el método HVSR de forma más rápida y accesible, mejorando la eficiencia de los estudios sísmicos y geotécnicos. Para ello, se sigue el procedimiento ilustrado en el diagrama de flujo de la Figura 1.



**Figura 1:** Diagrama de flujo de la metodología del trabajo.

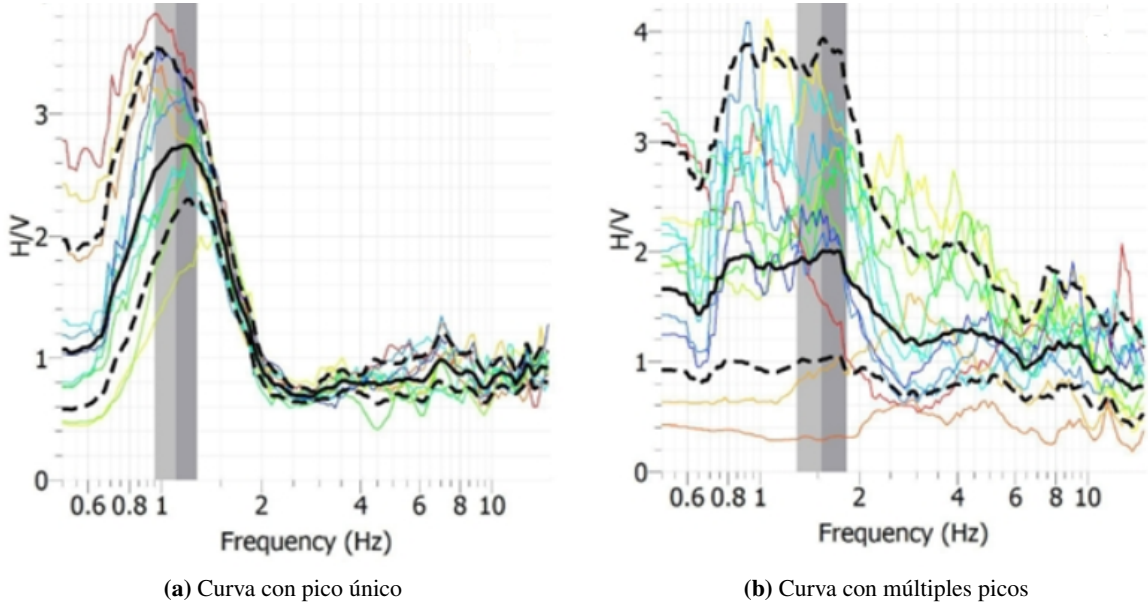
## 2.1. Recopilación de imágenes

Los datos son la materia prima para tareas de aprendizaje automático. En este caso, son imágenes HVSR generadas a partir del método de Nakamura, una técnica de sismica pasiva ampliamente utilizada para la caracterización del subsuelo. En total, se recopilaron 140 imágenes HVSR provenientes de artículos científicos, tesis de pregrado y doctorado, y pertenecientes al Dr. Olegario Alonso, docente de geofísica de la Universidad Central del Ecuador. Se detectó una alta variabilidad de formatos en este conjunto: algunas imágenes incluían ejes, otras trazos de colores y otras líneas entrecortadas asociadas a ruido. Para mitigar este problema y facilitar el aprendizaje de patrones visuales, se decidió emplear únicamente gráficas generadas con el programa geofísico *Geopsy*, que proporciona salidas uniformes, calidad gráfica y claridad espectral.

*Geopsy* es un conjunto de herramientas de código abierto diseñado para el procesamiento de datos de vibraciones ambientales obtenidos en campo, con aplicaciones en estudios de caracterización de sitios. Desde su primera versión en 2005, se ha consolidado como una herramienta fundamental en la geofísica aplicada, destacándose por su interfaz intuitiva y su disponibilidad gratuita [22]. Estas características lo hacen accesible tanto para profesionales como para

estudiantes, facilitando la interpretación de registros de vibración natural del terreno en distintos contextos de análisis sísmico y geotécnico.

Como resultado, 50 imágenes fueron descartadas mientras que 90 imágenes fueron consideradas válidas. Las imágenes seleccionadas comparten características homogéneas en cuanto a su formato y estructura, permitiendo identificar con claridad los parámetros espectrales clave: la amplitud máxima ( $A_0$ ) y la frecuencia fundamental ( $F_0$ ). Esta consistencia en el conjunto de datos es crucial para entrenar modelos de aprendizaje profundo con resultados confiables. En particular, se cuenta con imágenes provenientes de investigaciones [23][24] y proporcionadas por el Dr. Alonso, a partir de sus proyectos personales. La Figura 2a muestra un ejemplo de curva HVSR con un único pico predominante, lo cual facilita su interpretación automatizada. Por otro lado, la Figura 2b muestra una curva con múltiples picos, una situación más compleja que representa un reto tanto para los intérpretes humanos como para los modelos de IA, ya que dificulta la identificación precisa del pico fundamental.



**Figura 2:** Ejemplos de curvas HVSR.

Todas las imágenes fueron estandarizadas a un tamaño de 128×128 píxeles y almacenadas en formato JPG y PNG, utilizando servicios de almacenamiento en la nube como *Dropbox* y *Google Drive*. Esta estrategia facilitó el acceso compartido y la incorporación de nuevas imágenes desde distintos dispositivos, optimizando así el tiempo requerido para la recopilación de datos. A continuación, se procede a extraer la curva promedio de cada imagen, que servirá como entrada principal para el modelo de visión.

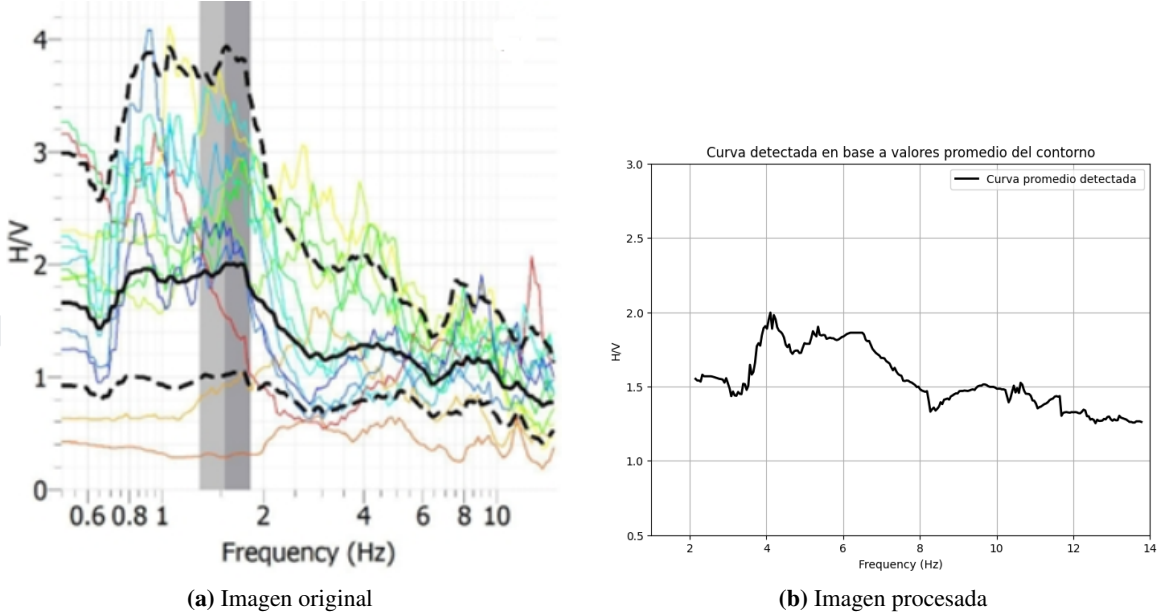
## 2.2. Extracción de curva promedio

Este paso significó uno de los principales desafíos del proyecto. Una imagen HVSR puede presentar múltiples trazos de distintos colores y grosores. A mayor cantidad de líneas superpuestas, mayor es la dificultad para que el modelo identifique la curva promedio relevante utilizada en la interpretación. Estas líneas adicionales fueron consideradas como “ruido”. La correcta identificación de la curva principal constituye la base para los cálculos posteriores de los parámetros  $F_0$  y  $A_0$ .

El preprocesamiento de las imágenes es fundamental para garantizar la calidad del dataset. Esta tarea se llevó a cabo mediante la librería *OpenCV* (Open Source Computer Vision Library) de Python en el entorno colaborativo *Google Colab*, lo que permitió una implementación ágil y replicable. Cada imagen es transformada a escala de grises (*cv2.cvtColor*), eliminando la información de color innecesaria y simplificando el análisis visual. Luego, se aplicó el algoritmo de detección de bordes *Canny* para identificar contornos relevantes, sobre todo aquellos asociados al trazo principal de la curva HVSR. Con estos bordes, se extrae la curva representativa de cada imagen, que contiene la información más relevante para estimar la frecuencia fundamental ( $F_0$ ) y su correspondiente amplitud ( $A_0$ ).

Además, se aplicaron técnicas de suavizado mediante filtros gaussianos para reducir el ruido visual, y se verificaron las dimensiones de cada imagen para asegurar compatibilidad con la arquitectura CNN propuesta. Estas operaciones permitieron generar un conjunto de datos limpio, estandarizado y listo para el etiquetado y posterior entrenamiento.

En la Figura 3, se presenta un ejemplo comparativo entre una imagen original generada en Geopsy y su versión procesada. Se observa que la imagen original (Figura 3a) contiene múltiples elementos visuales (líneas auxiliares, leyendas, ejes), mientras que la versión procesada (Figura 3b) resalta únicamente la curva promedio, lo cual mejora significativamente la interpretación automática.



**Figura 3:** Comparación entre la imagen sin procesar (izquierda) y la imagen luego del preprocesamiento (derecha), donde se ha aislado la curva promedio para su posterior análisis.

### 2.3. Etiquetado de imágenes

Una vez completada la recopilación y el preprocesamiento de las imágenes, estas fueron organizadas y almacenadas de forma estructurada. Dado que el modelo de aprendizaje automático supervisado requiere tanto los datos de entrada (imágenes HVSR) como los datos de salida (valores  $F_0$  y  $A_0$ ), se procede con el etiquetado manual de estas imágenes. Este proceso fue especialmente demandante, ya que implicó interpretar cada imagen de manera precisa. Se requirieron conocimientos técnicos en geofísica y del apoyo de herramientas especializadas para reducir la complejidad de la interpretación y minimizar errores.

El etiquetado consistió en identificar el pico más alto presente en cada imagen HVSR a partir del reconocimiento visual. Se estiman los valores de frecuencia fundamental del terreno ( $F_0$ ) y amplificación ( $A_0$ ), tomando en cuenta las escalas vertical y horizontal de la imagen. Así, se genera una primera interpretación, la cual se complementó con el cálculo del índice de vulnerabilidad ( $K_g$ ). Una descripción previa acompañó a cada imagen etiquetada, lo que fortaleció la calidad del conjunto de entrenamiento, tal como se indica en la Figura 4.

Cada nombre de imagen con sus correspondientes valores de frecuencia y amplitud se almacenaron en un archivo denominado *etiquetas.json*, guardado en Google Drive. Este archivo actúa como la referencia confiable (ground truth) para el futuro entrenamiento del modelo. A continuación, se muestra un ejemplo del formato de las etiquetas almacenadas:

```
{
  "imagen_001.png": {
    "frecuencia": 2.5,
    "amplitud": 5.0
  },
  "imagen_002.png": {
    "frecuencia": 1.70,
```

Código	interpretación	F0	A0	$K_g = \frac{A^2}{F_0}$
KVA_Pakistan_01	Se reconoce un pico (F0) en el basamento. La forma ancha del pico sugiere que el basamento tiene una inclinación. Su valor de índice de vulnerabilidad ( $K_g$ ) es de 6.44, esto sugiere que la zona no es vulnerable a daños o deformaciones en un evento sísmico.	1.2	2.78	6.4403
KVA_Pakistan_07	Es difícil identificar un pico representativo debido al comportamiento de la curva, sin embargo, se reconoce un pico amplio (F0) con bajas frecuencias lo que sugiere a que corresponde al basamento. La forma ancha del pico F0 sugiere que el basamento tiene una inclinación, su valor de índice de vulnerabilidad ( $K_g$ ) es de 2.22, esto sugiere que la zona es estable ante un evento sísmico. Posiblemente tenga poca cobertura de sedimentos.	1.8	2	2.2222

**Figura 4:** Ejemplo de interpretación y etiquetado de imágenes.

```

    "amplitud": 3.25
  }
}
```

Los pasos descritos han permitido generar un conjunto de datos robusto, coherente y con el nivel técnico necesario para entrenar y evaluar una CNN capaz de predecir con precisión los parámetros clave de una curva HVSR a partir de imágenes.

## 2.4. Arquitectura del sistema

Una vez que los datos se encuentran preparados, en esta etapa se establece cómo se procesarán para alcanzar el objetivo de automatizar la interpretación de imágenes HVSR. La estrategia consiste en combinar técnicas de visión por computadora y procesamiento del lenguaje natural, dando como resultado una arquitectura de tipo modular, la cual permite una organización clara y lógica del flujo de datos, facilitando tanto su implementación como su interpretación. El sistema propuesto integra tres módulos: visión, cálculo de vulnerabilidad y lenguaje; cada uno diseñado para una función específica en el proceso de análisis automatizado. A continuación, se detalla su funcionamiento.

### 2.4.1. Modelo de visión

Es la fase inicial que se encarga del procesamiento visual de la imagen espectral HVSR de entrada (Figura 5). Se aplica el preprocesamiento descrito en la Sección 2.2 para obtener la curva promedio. Una CNN que ha sido entrenada con el conjunto de imágenes ya etiquetadas, realiza la extracción de características para estimar los valores clave de esta curva, i.e. la frecuencia fundamental  $F_0$  y la amplitud máxima  $A_0$ . Estos dos parámetros son esenciales para la interpretación del comportamiento dinámico del subsuelo.

### 2.4.2. Cálculo de vulnerabilidad

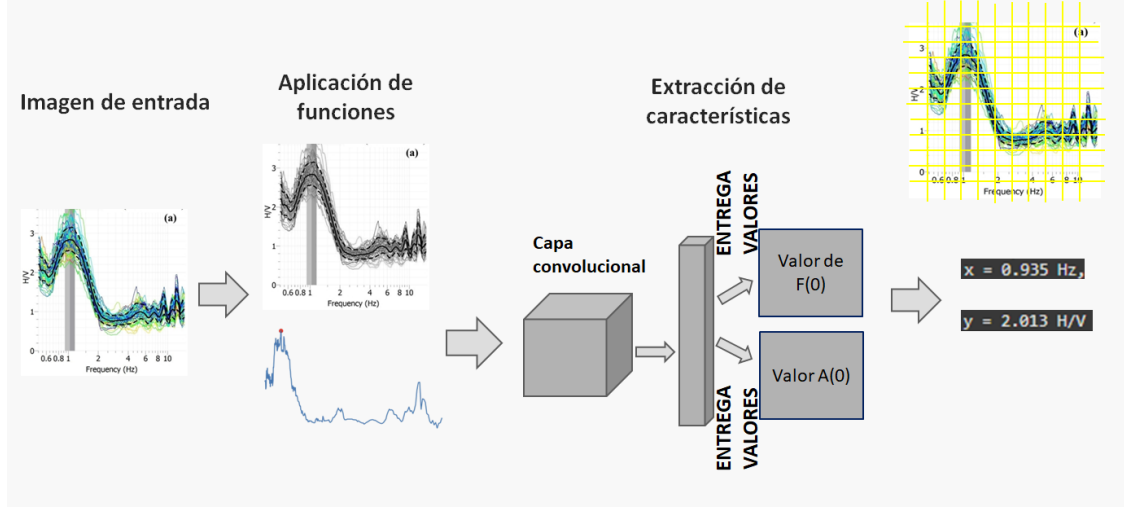
Una vez determinados  $F_0$  y  $A_0$ , se procede al cálculo del índice de vulnerabilidad sísmica  $K_g$ , utilizando estos valores obtenidos del módulo anterior. El índice se calcula mediante la siguiente fórmula:

$$K_g = \frac{A_0^2}{F_0} \quad (1)$$

Este índice representa una medida cuantitativa del riesgo o vulnerabilidad del sistema analizado, integrando variables físicas, geotécnicas y morfológicas. El resultado es un valor que puede ser interpretado en distintos niveles de criticidad, el cual permite evaluar la susceptibilidad del terreno a la amplificación de ondas sísmicas superficiales. Este valor proporciona una medida útil en contextos de microzonificación sísmica, estudios geotécnicos y evaluación del riesgo sísmico.

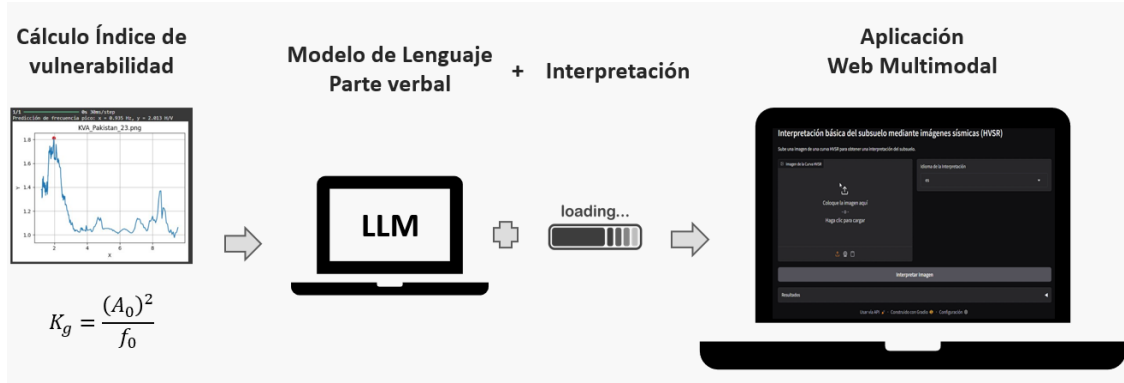
### 2.4.3. Modelo de lenguaje

El valor  $K_g$  calculado anteriormente es interpretado por un modelo de lenguaje, el cual se encarga de generar una descripción preliminar compuesta de datos numéricos y técnicos usando lenguaje natural. Esta descripción



**Figura 5:** Esquema de la arquitectura de la solución (Fase inicial).

contextualizada se presenta, tanto de manera textual como verbal, a través de una aplicación web, tal como se muestra en la Figura 6. Así, se facilita la comprensión para el usuario final y se promueve una toma de decisiones más informada.



**Figura 6:** Esquema de la arquitectura de la solución (Fase complementaria).

Este diseño en módulos ofrece una solución robusta e integradora, que no solo detecta patrones y calcula indicadores, sino que también comunica de manera efectiva sus resultados mediante el lenguaje humano. De esta manera, la arquitectura completa permite una transición fluida desde la entrada de datos visuales (imágenes HVSIR) hasta la salida de resultados cuantitativos e interpretativos que pueden ser utilizados por estudiantes y profesionales de la geociencia en contextos reales.

### 3. Experimentación y Resultados

Esta sección describe la experimentación llevada a cabo para validar la hipótesis de que, a partir de imágenes HVSIR preprocesadas, una CNN puede estimar con precisión la frecuencia fundamental  $F_0$  y la amplitud máxima  $A_0$  necesarias para el cálculo del índice de vulnerabilidad sísmica  $K_g$ . A continuación, se explica la plataforma computacional, el entrenamiento y la evaluación del modelo, y la extracción automática de los parámetros espectrales.

#### 3.1. Plataforma computacional

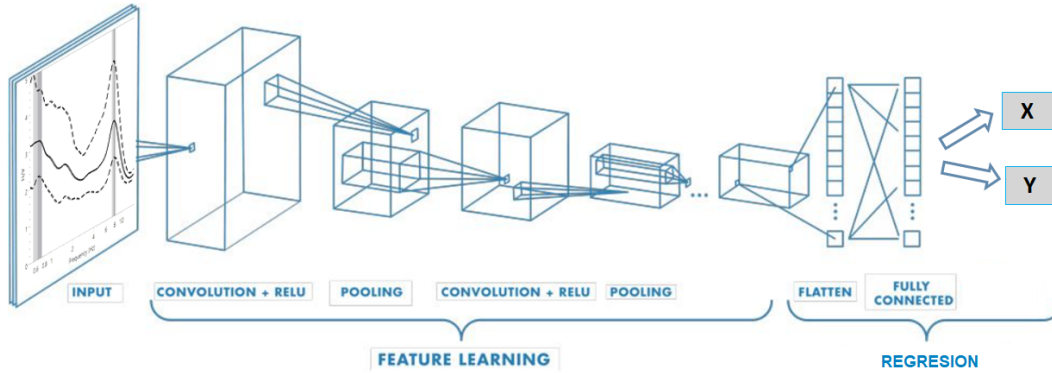
Inicialmente, se intentó desarrollar el proyecto de forma local en una computadora con procesador Intel Core i9-13900H a 2.60 GHz, 16 GB de RAM y sistema operativo Windows 11 Home de 64 bits. Sin embargo, se presentaron inconvenientes por falta de rendimiento, tiempo y almacenamiento. Finalmente, se optó por la plataforma *Google Colaboratory*, una herramienta gratuita en línea que proporciona un entorno colaborativo óptimo para el aprendizaje



y la ejecución en la nube [25]. Esta plataforma ofrece acceso a una GPU NVIDIA Tesla T4 de 16 GB de VRAM y 12.72 GB de RAM. El desarrollo utilizó el lenguaje de programación *Python 3.10* en un entorno de tipo *Jupyter Notebook*, con el apoyo de las siguientes librerías: *OpenCV* para el preprocesamiento de imágenes, *TensorFlow/Keras* para la implementación de la red neuronal y *NumPy*, *Pandas* y *Matplotlib* para el manejo, análisis y la visualización de datos. La integración con *Google Drive* aseguró la trazabilidad tanto de los datos como de los modelos generados.

### 3.2. CNN: configuración y entrenamiento

Una CNN es conocida por su eficacia comprobada en tareas de procesamiento de imágenes. Su estructura está inspirada en el sistema visual del ser humano y consta de dos partes: la extracción de características de la imagen de entrada (ojos de la red) y el cálculo de los valores (cerebro) a partir de dichas características. Para el desarrollo del modelo predictivo, se implementó una CNN cuya arquitectura se ilustra en la Figura 7.



**Figura 7:** Esquema de la arquitectura de la CNN utilizada. Modificado de MathWorks.

La red fue diseñada con una estructura sencilla pero funcional, compuesta por múltiples capas convolucionales con funciones de activación *ReLU*, seguidas de capas de *maxpooling* para reducir la dimensionalidad y evitar el sobreajuste. Este mecanismo permite extraer características espaciales relevantes de las imágenes HVSR de entrada que son fundamentales para estimar los parámetros físicos del subsuelo. Luego, las salidas de estas capas se convierten en un vector de características mediante la capa de *flatten*, y que se pasa a capas densas para culminar en una capa de salida con dos neuronas, las cuales predicen los valores continuos ( $X, Y$ ); es decir, la frecuencia fundamental  $F_0$  y la amplitud máxima  $A_0$ , respectivamente. Se utilizó la función de pérdida *MSE* (Mean Squared Error), ideal para problemas de regresión, y el optimizador *Adam*, conocido por su eficiencia en la convergencia del entrenamiento.

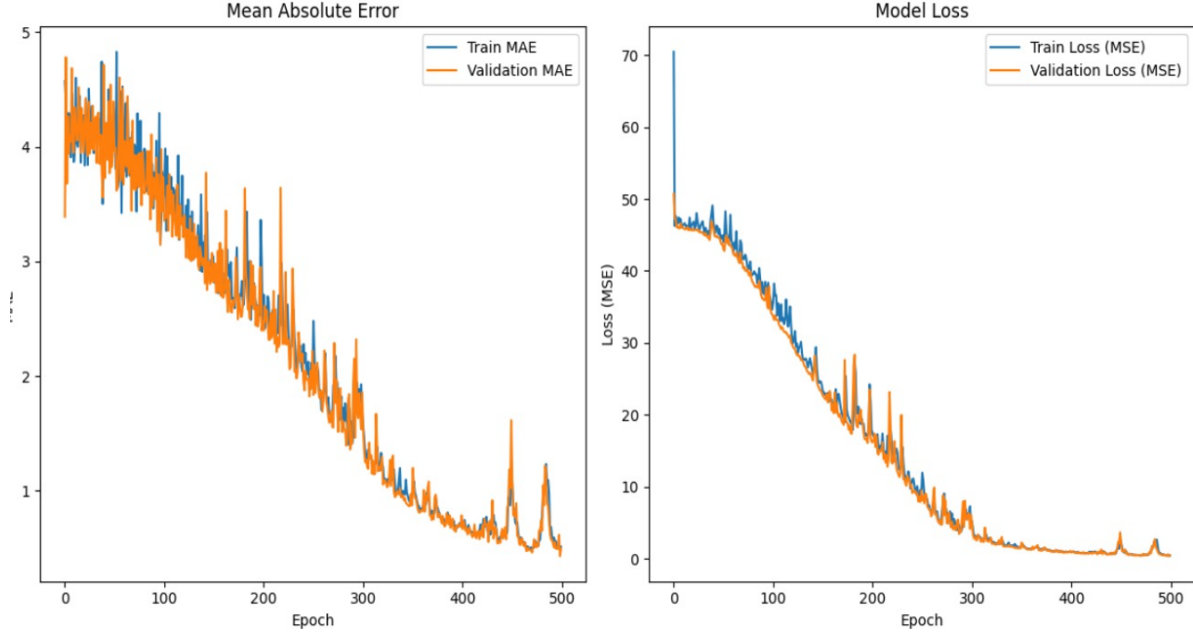
Previo al entrenamiento, el conjunto de imágenes se dividió en un 80 % para entrenamiento y validación, y un 20 % para prueba. El modelo fue entrenado durante 500 épocas, monitoreando las métricas de Error Cuadrático Medio (MSE) y el Error Absoluto Medio (MAE). La Figura 8 presenta las curvas de aprendizaje del modelo para el MAE y el MSE durante el entrenamiento. Se observa una tendencia decreciente en ambas métricas, tanto en el conjunto de datos de entrenamiento como en el de validación. La convergencia de ambas curvas hacia valores próximos a 0, que indica un comportamiento ideal, sin evidencias de sobreajuste, y demuestra que el modelo logró aprender los patrones subyacentes de los datos de entrenamiento.

### 3.3. Evaluación del modelo

Para evaluar la capacidad predictiva del modelo, se utilizó el conjunto de prueba compuesto por el 20 % de las imágenes, que no fueron vistas durante el entrenamiento. De esta manera, se garantizó que el modelo se enfrentara a imágenes completamente nuevas, lo que permitió obtener una evaluación objetiva de su rendimiento. Se utilizan las métricas de error de regresión MAE y MSE, que son apropiadas para evaluar la precisión de un modelo que predice valores continuos. El MAE mide la desviación promedio de los valores predichos del modelo con respecto a los valores reales de los parámetros de frecuencia y amplitud [26]. El MSE, al elevar al cuadrado la diferencia entre los valores predichos y reales, penaliza más fuertemente los errores grandes. Esto es particularmente útil para identificar si el modelo tiene dificultades con algún valor atípico o extremo en los datos [27].

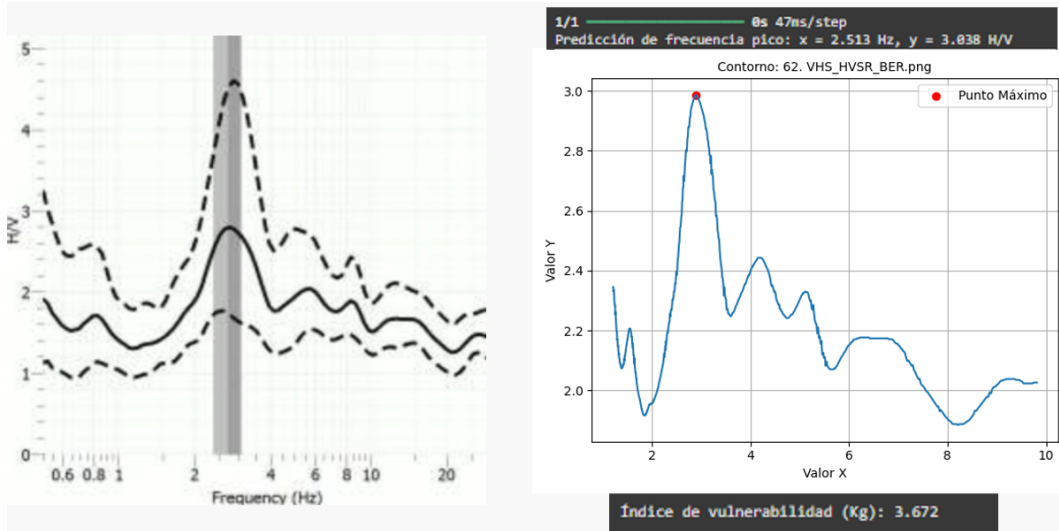
La correlación entre los valores reales y predichos de  $F_0$  y  $A_0$  fue adecuada, con desviaciones mínimas en los casos más complejos. Esto demuestra que el modelo es capaz de predecir los parámetros geofísicos con una alta precisión en imágenes que nunca antes había visto. Por ende, el desempeño del modelo sugiere una generalización adecuada.





**Figura 8:** Curvas de aprendizaje del modelo: evolución del MAE y MSE durante el entrenamiento.

Una vez que se ha validado el buen desempeño del modelo, se procede a automatizar la obtención de los parámetros espectrales clave. A partir de la imagen HVSR procesada, se generó una máscara binaria, de la cual se aisló la curva promedio dominante. Mediante la detección de contornos, se identificó el pico principal de la curva, obteniendo así  $F_0$  y  $A_0$ . Con estos valores se calculó el índice de vulnerabilidad sísmica utilizando la Ecuación 1, tal como se muestra en la Figura 9.



**Figura 9:** Filtrado de la curva dominante: imagen HVSR con múltiples trazos (izquierda) y curva promedio aislada empleada para estimar  $F_0$  y  $A_0$  (derecha).

El flujo de trabajo descrito evidencia que una CNN puede estimar con precisión los parámetros espectrales fundamentales de una curva HVSR a partir de imágenes preprocesadas, sin intervención humana. Esta estimación automatizada permite calcular de forma directa el índice de vulnerabilidad sísmica  $K_g$ , el cual se convierte en el insumo principal para producir interpretaciones básicas del subsuelo utilizando un LLM. Esta experimentación se detalla a continuación.

### 3.4. Modelo de Lenguaje Natural

La solución propuesta busca complementar la visión computacional basada en una CNN con el procesamiento del lenguaje natural (NLP) usando un modelo de lenguaje a gran escala (LLM). Por un lado, la CNN se encarga de la extracción automática de la frecuencia pico y la amplitud máxima a partir de imágenes sísmicas HVSR; por otro lado, el LLM tiene la función de interpretar el índice de vulnerabilidad calculado con estos valores y generar una descripción textual del subsuelo que sea accesible, contextualizada y significativa para el usuario final.

A diferencia de la CNN, que pudo ser entrenada sin problemas, el reentrenamiento de un LLM representa un desafío y es prácticamente inviable en entornos como la versión gratuita de Google Colab y en equipos locales, debido a los altos requisitos de memoria RAM y GPU. Como alternativa, se opta por “instruir” al modelo aprovechando el vasto conocimiento del LLM, en lugar de “enseñar” al modelo. A través de un *prompt*, es posible guiar el comportamiento del LLM mediante instrucciones estructuradas cuidadosamente diseñadas para alcanzar los resultados deseados [28]. Para facilitar la interpretación de este índice, se diseña un *prompt* estructurado en lenguaje natural que permite traducir los resultados técnicos en una interpretación comprensible del subsuelo.

#### 3.4.1. Diseño del prompt

Este proceso es especialmente útil cuando se desea obtener salidas textuales específicas y adaptadas a un determinado formato, como lo exige una descripción relacionada con la caracterización técnica del subsuelo. La construcción del *prompt* fue realizada de manera iterativa, refinando una versión inicial breve y general hasta lograr una versión más elaborada y específica, la cual permite obtener una respuesta del LLM alineada con la del experto.

El *prompt* se basa en una plantilla de texto que tiene la forma: *instrucción + entrada*; es decir, una frase que simula una pregunta o indicación clara y el valor del índice de vulnerabilidad obtenido, el cual es insertado automáticamente en la plantilla del *prompt*. Por ejemplo:

“Genera una interpretación del índice de vulnerabilidad sísmica de 4.6...”

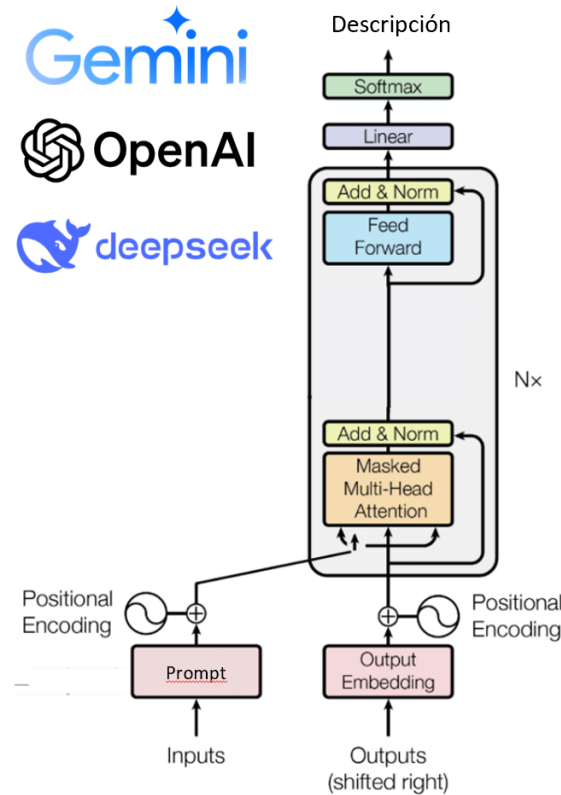
Este paso es crucial, ya que una formulación ambigua o incorrecta del *prompt* puede inducir al LLM a generar respuestas erróneas o incoherentes, fenómeno conocido como *alucinación* [29]. Con el diseño del *prompt* ya establecido, el siguiente paso es la ejecución de la consulta. A continuación, se detalla cómo el LLM utiliza estas instrucciones para producir la descripción textual de salida.

#### 3.4.2. Ejecución del LLM

La descripción automática del subsuelo se logra una vez que el LLM recibe el *prompt* completo, lo procesa y genera una respuesta textual contextualizada. La capacidad del LLM para traducir variables numéricas en explicaciones comprensibles es una de sus funcionalidades más valiosas [30]. Esta salida puede ser, por ejemplo:

“El índice de vulnerabilidad sísmica calculado es de 4.6, lo que indica un nivel moderado de amplificación del suelo...”

Para generar esta respuesta, el funcionamiento del LLM se basa en un modelo *Transformer* de tipo *decoder-only* [31]. En este diseño no hay codificador separado, ya que todo el *prompt* constituye una sola secuencia y el modelo aprende a predecir la siguiente palabra condicionada por todas las palabras anteriores. Esto ha demostrado ser altamente efectivo en tareas de generación y comprensión de texto. La Figura 10 ilustra el flujo de procesamiento del LLM a través de una estructura jerárquica de varias capas que permiten la generación de la descripción en la fase de inferencia.



**Figura 10:** Esquema del funcionamiento del modelo de lenguaje natural.

1. El modelo recibe el prompt para ejecutar la consulta. Este texto de entrada se descompone en una secuencia de palabras (*tokens*) que son procesadas de forma simultánea.
2. Cada token se convierte en un vector numérico o *embedding*, que representa el concepto de cada palabra. Estos vectores residen en un espacio de alta dimensionalidad, en el cual la proximidad refleja similitud semántica, mientras que la lejanía indica significados distintos.
3. La codificación posicional inyecta información al embedding sobre la posición de cada token en la secuencia, lo que permite al modelo “saber” el orden de las palabras, mantener el sentido gramatical y la estructura del texto de entrada.
4. El mecanismo de autoatención enmascarada hace que cada token obtenga una representación contextualizada a partir de los tokens anteriores en la secuencia. Para lograrlo, cada token “consulta” a los tokens que lo preceden y a sí mismo, asignando un peso de influencia según su relevancia. Con estos pesos se calcula una combinación ponderada de los vectores, produciendo así un vector de contexto para cada token. Este resultado representa el significado del token no solo de forma aislada, sino también en relación con el contexto previo.
5. Una red neuronal feed-forward incorpora parámetros adicionales y funciones de activación no lineales, lo que añade flexibilidad al modelo. Esto permite capturar dependencias complejas en los datos y producir representaciones más refinadas de cada token.
6. Por último, el modelo se enfoca en el vector del último token, el cual incorpora información de toda la secuencia previa. Una transformación lineal proyecta este vector al espacio del vocabulario. Luego, una función *softmax* produce una distribución de probabilidad del siguiente token, que se convierte en la respuesta del LLM.

Los pasos descritos se efectúan de manera secuencial y autoregresiva; es decir, una vez que se predice el primer token de la respuesta, este se añade al final de la secuencia de entrada, y el proceso se repite para predecir el siguiente token. Esto es lo que permite que el modelo genere texto detallado de forma coherente y contextualizada, actuando como una capa semántica de interpretación, lo que es fundamental en la humanización de los resultados: convertir cifras abstractas en explicaciones relevantes, alineadas con el conocimiento práctico de los usuarios.

## 4. Aplicación web multimodal

Nuestro objetivo es proporcionar una herramienta práctica y accesible que automatiza la interpretación básica de curvas HVSR, útil para estudios geotécnicos y sísmicos, entre otros. Esta sección expone el desarrollo de una aplicación web multimodal que aprovecha la salida generada por el LLM, presentándola de forma legible para el usuario, acompañada de opciones de idioma, accesibilidad mediante síntesis de voz y elementos visuales de apoyo como gráficos de dispersión, lo que puede adaptarse a distintos públicos y necesidades, mejorando la experiencia general del usuario.

Con el propósito de democratizar el acceso a la herramienta y fomentar la posibilidad de replicar el estudio, la aplicación se encuentra publicada y puede ser desplegada a través de *Hugging Face*<sup>1</sup>, que es una plataforma abierta reconocida por su ecosistema de modelos con IA y herramientas colaborativas [32] que proporciona una base estable y escalable, permitiendo a estudiantes, investigadores y profesionales interactuar con el modelo de forma libre y transparente.

El código y la aplicación están disponibles en el repositorio del proyecto con el siguiente enlace: <https://tinyurl.com/IAInterpretacionBasicaHVSr>. Los pasos necesarios para la correcta carga, documentación y puesta en funcionamiento de la aplicación alojada en esta plataforma son:

1. *Preparación del modelo*: el modelo CNN previamente entrenado y guardado en un formato compatible, como .h5 (Keras), .pt (PyTorch) o .pkl (Pickle). Además, es fundamental disponer del código necesario para cargar el modelo y realizar inferencias, usualmente escrito en Python, así como cualquier archivo auxiliar (por ejemplo, etiquetas, funciones de preprocesamiento o scripts de visualización).
2. *Creación de un repositorio en Hugging Face*: Es necesario contar con una cuenta en Hugging Face. Una vez iniciada la sesión, se debe crear un nuevo repositorio en la sección **Models** para compartir el modelo como recurso.
3. *Elección del framework*: Hugging Face permite seleccionar el framework de implementación. Gradio es recomendable para modelos de visión por computadora por su facilidad para crear interfaces amigables. Para su implementación se empleó la biblioteca *Gradio* [33], un framework de código abierto que facilita el desarrollo rápido de interfaces para modelos de Machine Learning. Su integración en entornos como Google Colab permite un prototipado ágil y una depuración eficiente.
4. *Despliegue*: en este espacio se deben incluir los archivos principales:
  - `app.py` (archivo con la lógica de la interfaz)
  - `requirements.txt` (librerías necesarias, como `tensorflow`, `opencv-python`, `gradio`, etc.)
  - El archivo del modelo (`modelo.h5` u otro formato)
  - Otros archivos de configuración de ser necesario.

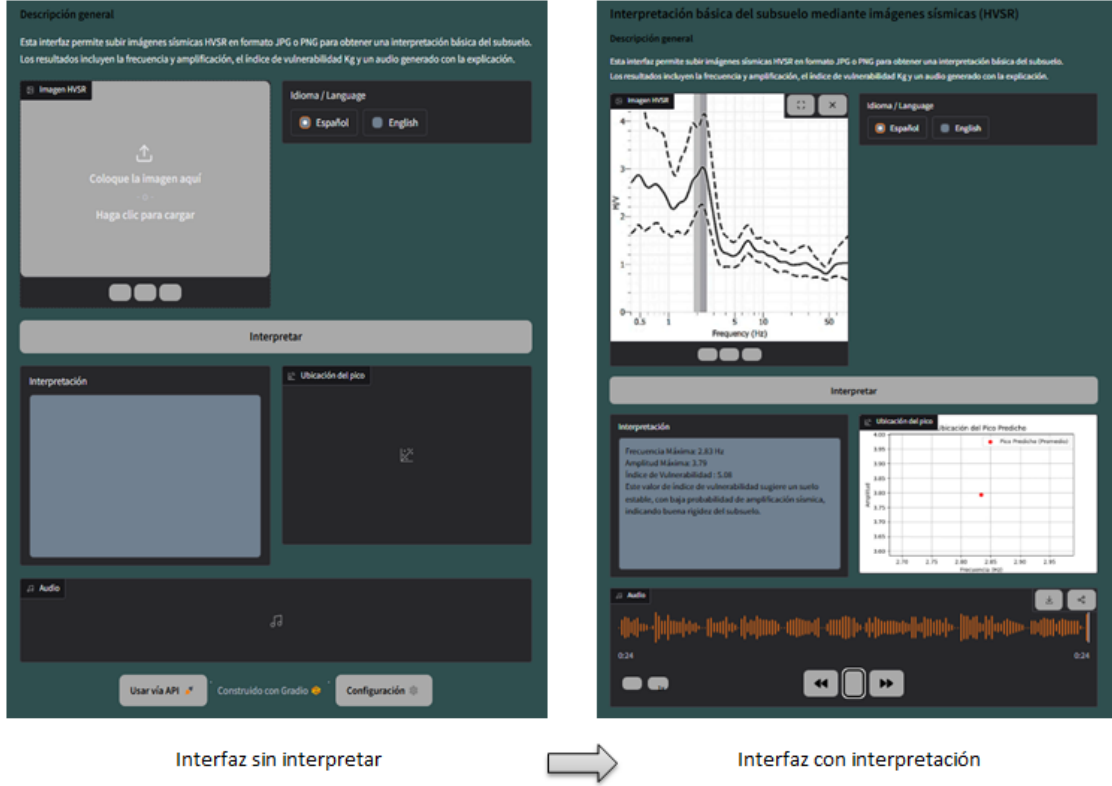
Finalmente, el espacio de Hugging Face al detectar los archivos `app.py`, `modelo.h5` y `requirements.txt`, instala automáticamente las dependencias y lanza la aplicación con la interfaz definida, tal como muestran las capturas de pantalla de la Figura 11.

El uso y funcionamiento del sistema sigue un flujo de entrada, proceso y salida. Esta aplicación se despliega a través de una interfaz interactiva que le permite al usuario cargar una imagen sísmica HVSR para ser interpretada. La imagen es procesada por el modelo de red convolucional que extrae las características numéricas clave (frecuencia y amplitud). Estos valores se utilizan para calcular el índice de vulnerabilidad ( $K_g$ ) y generar una interpretación textual mediante el modelo de lenguaje. Esta respuesta presenta la información cuantitativa derivada del análisis (frecuencia, amplitud e índice  $K_g$ ) acompañada de una descripción cualitativa del subsuelo. Esta descripción, generada mediante *prompt engineering* aplicado a un modelo de lenguaje, contextualiza los resultados en términos comprensibles. Por ejemplo:

“... sugiere un suelo estable, con baja probabilidad de amplificación sísmica ...”

La salida es de tipo multimodal: textual, gráfica y auditiva. Además, la interfaz permite seleccionar el idioma de la interpretación (español o inglés), aumentando la versatilidad de la herramienta. La modalidad gráfica consiste en un diagrama de dispersión bidimensional que posiciona el punto correspondiente a los valores predichos por el modelo (frecuencia y amplitud máxima). Esta visualización facilita la comprensión de la ubicación relativa de la muestra respecto al espacio de características del conjunto de entrenamiento, brindando una perspectiva espacial de la interpretación. La modalidad de audio se produce mediante un motor de síntesis de voz (*TTS*), que convierte la salida textual en un archivo de audio. Esta funcionalidad no solo mejora la accesibilidad, sino que también permite al usuario recibir los resultados en el idioma elegido, con una entonación natural que favorece la comprensión auditiva del contenido.

<sup>1</sup><https://huggingface.co>



**Figura 11:** Capturas de pantalla de la aplicación web multimodal creada con Gradio.

Cabe destacar que, durante la etapa de entrenamiento del modelo convolucional, se trabajó con la curva promedio completa obtenida a partir del procesamiento de imágenes HVSR mediante detección de contornos. Esta representación permitió identificar con precisión el pico fundamental de la curva ( $F_0$ ) y su amplitud asociada ( $A_0$ ), parámetros esenciales para el cálculo del índice de vulnerabilidad sísmica. Sin embargo, en la fase de implementación de la interfaz web, se optó por simplificar la visualización final mostrando únicamente el punto correspondiente al valor máximo estimado por el modelo en el espacio de frecuencia y amplitud como respuesta a consideraciones prácticas relacionadas con la optimización del desempeño de la aplicación en entornos web, especialmente en plataformas gratuitas de despliegue como *Hugging Face Spaces*, las cuales presentan limitaciones en términos de recursos computacionales. Adicionalmente, la renderización de la curva promedio completa implicaría una mayor carga de datos en el navegador del usuario, lo que podría afectar la experiencia en dispositivos móviles o con capacidad de procesamiento reducida. Por lo tanto, se priorizó una visualización ligera y funcional, que mantenga la interpretación sísmica esencial sin comprometer el rendimiento ni la accesibilidad de la aplicación. Además, se debe considerar que las imágenes cargadas en la página web deben tener un formato similar al de las imágenes de prueba disponibles en el enlace proporcionado al final del trabajo, con el fin de asegurar una interpretación lo más precisa posible. Se espera un rendimiento óptimo del sistema cuando las características de las imágenes son similares a las del conjunto de entrenamiento, lo que asegura resultados fiables.

## 5. Conclusiones y trabajo futuro

Se desarrolló una solución basada en IA para la automatización del análisis de imágenes HVSR, capaz de replicar con alto grado de precisión la interpretación realizada por un especialista en geofísica. El flujo de trabajo propuesto integra procesamiento de imágenes, predicción automática de la frecuencia fundamental ( $F_0$ ) y la amplitud máxima ( $A_0$ ), cálculo del índice de vulnerabilidad sísmica ( $K_g$ ) y procesamiento de lenguaje natural, lo que permite reducir significativamente el tiempo y la subjetividad asociados al análisis tradicional.

El entrenamiento del modelo de visión reflejó un aprendizaje progresivo y una convergencia adecuada, con errores bajos de las métricas de regresión (MAE y MSE) y sin evidencia de sobreajuste. El modelo entrenado mostró buena

capacidad de generalización frente a imágenes no vistas. La estandarización del dataset mediante el uso exclusivo de imágenes generadas por Geopsy contribuyó significativamente a la estabilidad y robustez del modelo.

Los resultados obtenidos mostraron una alta coincidencia entre los valores predichos por el modelo y aquellos determinados por geofísicos expertos, lo que respalda la validez de la propuesta. Por ende, esta metodología puede ser aplicada como apoyo técnico en investigaciones geofísicas o estudios de microzonificación sísmica, especialmente cuando se requiere interpretar grandes volúmenes de datos en tiempos reducidos, y la posibilidad de integrar este enfoque en herramientas más amplias de evaluación geotécnica. La solución propuesta abre nuevas oportunidades para la incorporación de IA en la geología aplicada, haciendo más accesible el análisis espectral a usuarios no especializados.

El éxito de este proyecto abre para futuros desarrollos y mejoras. Se puede avanzar en la integración de parámetros geofísicos adicionales, como la velocidad de las ondas de corte (VS), para enriquecer la interpretación. Uno de los principales desafíos es la heterogeneidad en los formatos de las imágenes HVSR, las cuales presentaban ejes, trazos de colores o líneas asociadas a ruido, dificultando el aprendizaje. Del mismo modo, la arquitectura del modelo podría adaptarse para analizar múltiples picos en un mismo espectro, ofreciendo así una interpretación más detallada de la estratigrafía del subsuelo. Estas mejoras directas permitirían que la herramienta sea aún más robusta y completa para su aplicación en estudios.

## Disponibilidad de datos y código

Los datos y el código que respaldan los hallazgos del presente estudio están disponibles en los siguientes enlaces:

Aplicación web: <https://tinyurl.com/IAInterpretacionBasicaHVSR>

Conjunto de imágenes de prueba\*: <https://tinyurl.com/PublicImagesHVSR>

\* Por razones de confidencialidad y protección de datos, no es posible compartir la totalidad del conjunto de imágenes HVSR utilizadas en este estudio. Sin embargo, se ha habilitado una carpeta con una muestra representativa y limitada de imágenes, suficiente para probar y utilizar adecuadamente el software desarrollado.

## Agradecimientos

Los autores expresan su más sincero agradecimiento al Ingeniero Alonso Olegario Pandaneves, docente de la Facultad de Ingeniería en Geología, Minas, Petróleos y Ambiental de la Universidad Central del Ecuador, por su invaluable guía, conocimientos compartidos y apoyo constante durante el desarrollo de este proyecto. Su experiencia en geofísica aplicada y métodos de interpretación sísmica fue fundamental para la correcta orientación del trabajo, aportando una visión técnica y académica de gran valor.

## Referencias

- [1] C. Merino. Estudios de microzonificación sísmica, de interacción suelo estructura y licuefacción. *Ingeniería Civil*, 2019. <https://ingenieriacivil.cedex.es/index.php/ingenieria-civil/article/view/2400>.
- [2] O. M. Alonso Pandaveles. *Técnicas de sísmica pasiva HVSR aplicadas a la geotecnia. Aplicación al estudio de movimientos en masa en la planificación territorial e infraestructura civil en Ecuador*. PhD thesis, Universitat Politècnica de València, 2024.
- [3] M. Miezah Adams, F. K. Torvor, E. Ansah, E. Boateng, and A. Ewusi. Microtremor hvsr technique for seismic risk vulnerability studies and microzonation of site materials. *Geophysical Research Journal*, 2024.
- [4] Mohamed A Gamal, Mohamed H Khalil, and George Maher. Monitoring and studying audible sounds inside different types of soil and great expectations for its future applications. *Pure and Applied Geophysics*, 177(11):5397–5416, 2020.
- [5] Y. Nakamura. A method for dynamic characteristics estimation of subsurface using microtremors on the ground surface. *Railway Technical Research Institute, Quarterly Reports*, 30(1):25–33, 1989.
- [6] Swathi Priyadarsini Putti and Neelima Satyam. Evaluation of site effects using hvsr microtremor measurements in vishakhapatnam (india). *Earth Systems and Environment*, 4(2):439–454, 2020.
- [7] E. Giolo, F. Calderón, S. Pagano, S. Via Gatica, and C. Frau. La técnica del cociente espectral h/v como herramienta para la determinación de estructuras ciegas. un caso de estudio en la ciudad de mendoza, argentina. In *ASAGI*, Mendoza, 2019.

- [8] F. Calderón. Utilización del método hvsr para la determinación de estructuras geológicas ciegas. Technical report, International Society for Soil Mechanics and Geotechnical Engineering (ISSMGE), 2024. MT 4 – FV (Proceeding FV 186).
- [9] O. Alonso-Pandavenes, F. J. Torrijo Echarri, and J. Garzón-Roca. Gestión sostenible de deslizamientos de tierra en ecuador: Aprovechamiento de estudios geofísicos para la reducción eficaz del riesgo. *Sustainability*, 16(24):10797, 2024.
- [10] Marco Mucciarelli and Maria Rita Gallipoli. A critical review of hvsr technique from ambient noise analysis: its advantages and limitations. *Journal of Earthquake Engineering*, 5(Special Issue):1–25, 2001.
- [11] M. Faris, S. Kurniawan, B. Pranata, and S. Prayitno. Seismic microzonation based on hvsr inversion results for shear wave (vs30) mapping and soil vulnerability in west sulawesi and south sulawesi regions. *Journal of Seismology and Earthquake Engineering*, 2023. In press or journal not specified, complete reference needed for formal use.
- [12] U.S. Environmental Protection Agency. Passive seismic (hvsr). <https://www.epa.gov/environmental-geophysics/passive-seismic-hvsr>, 2025. Accessed January 6, 2025.
- [13] Y. Nakamura. On the h/v spectrum. In *The 14th World Conference on Earthquake Engineering*, pages 1–10, 2008.
- [14] V. Salinas. *Detección fina de cambios laterales en geología superficial y en los suelos y caracterización de su influencia en la respuesta sísmica local*. PhD thesis, Universitat Politècnica de Catalunya, Barcelona, España, 2015. <https://www.tdx.cat/bitstream/handle/10803/334695/TVSN4de4.pdf?sequence=15.xml>.
- [15] Iván Lizarazo. *Geografía y cultura: territorio, paisaje y representaciones espaciales*. Universidad Nacional de Colombia, Bogotá, Colombia, 2008.
- [16] P. Joseph. Autohvsr: Automatización del procesamiento de la técnica hvsr mediante inteligencia artificial. *HVSR Web - Software Documentation*, 2023. Disponible en: <https://hvsrweb.com>.
- [17] E. Zurita Romero. Determinación de facies a partir de registros geofísicos de pozo y propiedades petrofísicas usando herramientas de inteligencia artificial. Tesis de maestría, Benemérita Universidad Autónoma de Puebla, 2020. Accedido el 6 de agosto de 2025.
- [18] W.A. Faris and A.N. Othman. Prediction of soil shear strength using artificial neural networks (ann): A comparative study. *International Journal of Advanced Computer Science and Applications*, 14(2):609–616, 2023.
- [19] José Pérez Esquivel, Marco Pérez, and Cristina Juárez Landin. Aprendizaje de las redes neuronales convolucionales mediante el análisis y modificación de la clasificación de imágenes para mejorar su precisión, 08 2024.
- [20] Haotian Liu, Chunyuan Li, Qingyang Wu, and Yong Jae Lee. Visual instruction tuning, 2023.
- [21] Jiquan Ngiam, Aditya Khosla, Mingyu Kim, Juhan Nam, Honglak Lee, and Andrew Y Ng. Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)*, pages 689–696, 2011.
- [22] Marc Wathelet, Jean-Laurent Chatelain, Cécile Cornou, Giuseppe Di Giulio, Bertrand Guillier, Matthias Ohrnberger, and Alexandros Savvaidis. Geopsy: A User-Friendly Open-Source Tool Set for Ambient Vibration Processing. *Seismological Research Letters*, 91(3):1878–1889, 2020.
- [23] M. Rivera, A. Ceballos, J. Sánchez, A. Fuentes, and J. Pardo. Directional hvsr at the chalco lakedeb zone of the valley of méxico: Analysis and interpretation. *Elsevier*, 2024.
- [24] Giancarlo Dal Moro. Insights on surface wave dispersion and hvsr: Joint analysis via pareto optimality. *Elsevier*, 2010.
- [25] Douglas Viera. Google colab: ¿qué es y cómo usarlo?, 2024. <https://www.hostgator.net.ec/blog/google-colab/>.
- [26] C. J. Willmott and K. Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate Research*, 30:79–82, 2005.
- [27] T. Chai and R. R. Draxler. Root mean square error (rmse) or mean absolute error (mae)? – arguments against avoiding rmse in the literature. *Geoscientific Model Development*, 7:1247–1250, 2014.
- [28] Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed H Chi, Quoc V Le, and Denny Zhou. Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*, 2022.
- [29] Laria Reynolds and Kyle McDonell. Prompt programming for large language models: Beyond the few-shot paradigm. *arXiv preprint arXiv:2102.07350*, 2021.
- [30] Tom B Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.



- [31] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [32] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, and Alexander M Rush. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pages 38–45. Association for Computational Linguistics, 2020.
- [33] Abubakar Abid, Maheen Abdalla, and James Zou. Gradio: Hassle-free sharing and testing of ml models in the wild. *arXiv preprint arXiv:1906.02569*, 2019.
- Viera, Douglas. *Google Colab: ¿Qué es y cómo usarlo?* 2024. Disponible en: <https://www.hostgator.net.ec/blog/google-colab/>.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., ... & Le, Q. (2022). Emergent abilities of large language models. *arXiv preprint arXiv:2206.07682*.
- Reynolds, L., & McDonell, K. (2021). Prompt Programming for Large Language Models: Beyond the Few-Shot Paradigm. *arXiv preprint arXiv:2102.07350*.
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in neural information processing systems*, 33, 1877–1901.
- Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., & Ng, A. Y. (2011). Multimodal deep learning. In *Proceedings of the 28th international conference on machine learning (ICML-11)* (pp. 689–696).
- Abid, A., Abdalla, M., & Zou, J. (2019). Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild. *arXiv preprint arXiv:1906.02569*. <https://doi.org/10.48550/arXiv.1906.02569>
- Baltrušaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal Machine Learning: A Survey and Taxonomy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 423–443. <https://doi.org/10.1109/TPAMI.2018.2798607>
- Schroeder, D. (2021). Text-to-Speech Synthesis and Its Applications in Accessibility Technologies. *International Journal of Human-Computer Interaction*, 37(11), 1083–1095. <https://doi.org/10.1080/10447318.2021.1917884>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., ... & Rush, A. M. (2020). Transformers: State-of-the-Art Natural Language Processing. En *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 38–45). <https://doi.org/10.18653/v1/2020.emnlp-demos.6>

## Autores

### Renata Paola Bohórquez Kon



Nació en Quito el 10 de septiembre de 1998, pero creció en la ciudad de Esmeraldas. Cursa 10mo semestre de la carrera de Geología en la Universidad Central del Ecuador. Fue pasante de geociencias en PCR. Posee experiencia en geología del subsuelo, geología del petróleo, aplicación de sismica pasiva y lenguaje Python. Ganadora del segundo lugar en el Imperial Barrel

Award (IBA) 2025 en la Región América Latina y el Caribe.

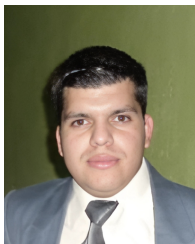
### Katherine Vanessa Ayala Soatunce



Nació en La Maná, el 3 de mayo de 1993. Cursa 10mo semestre de la carrera de Geología en la Universidad Central del Ecuador. Pasante del Centro de Investigaciones Quito (CIQ) EP Petroecuador. Tiene experiencia en Cartografía geológica, AutoCAD, SIG, RecMin, Teledetección, Geología del subsuelo, Gestión de Riesgos y desarrollo e implementación de modelos de

aprendizaje profundo utilizando redes neuronales y herramientas como Google Colaboratory.

### Víctor Hugo Solís Acosta



Nació en Quito el 28 de febrero de 1990. Cursa 10mo semestre de la carrera de Geología en la Universidad Central del Ecuador. Fue pasante de geociencias en PETROECUADOR EP. Posee experiencia laboral de 10 años como dibujante de planos AS BUILT, topografía, asistente de campo en geofísica de refracción y geotecnia, además de co-

nocimiento en manejo de softwares como: ArcGIS, AutoCAD 2D y 3D, RecMin, Google Colab, entre otros.

### Eduardo Gabriel Guerrero Fuel



Nacido en Tulcán, Carchi, Ecuador, el 28 de septiembre de 1995, estudiante de 10mo semestre de Ingeniería en Geología en la Universidad Central del Ecuador. Experiencia en el sector privado como operador de dron, manejando SIG, mapeo y teledetección. Así como en la industria pública como interno en empresas petroleras, en áreas de explotación y desarrollo, utilizando registros eléctricos.

### Christian Iván 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.