

Diseño e implementación de una red neuronal convolucional para la estimación de mecanismos focales a partir de datos sismológicos sintéticos basados en la región del Valle Medio del Magdalena

CRISTIAN ALBERTO PEÑA CALA¹, SERGIO ANDRES GIL MORENO²,
JHEYSTON OMAR SERRANO LUNA³, AND ANA BEATRIZ RAMIREZ SILVA⁴

¹Escuela de Ingeniería Eléctrica, Electrónica y Telecomunicaciones (E3T), 2023

²Universidad Industrial de Santander (UIS), 2023

³Connectivity and Signal Processing (CPS), 2023

⁴cps.info@uis.edu.co

Compilado Julio 9, 2023

En este artículo se presenta el diseño e implementación de un Modelo de Red Neuronal Convolucional (CNN) para estimar los mecanismos focales basado en la forma de onda de mediciones de velocidad sísmica. Este enfoque se propone como una alternativa novedosa a los métodos tradicionales de estimación de mecanismos focales. La aplicación de inteligencia artificial y aprendizaje automático ofrece nuevas oportunidades y contribuciones significativas en el campo de la sismología. En particular, las CNN son capaces de extraer características relevantes y reconocer patrones complejos en los datos sísmicos, lo que las convierte en una herramienta prometedora para la estimación de mecanismos focales. En el entrenamiento de la CNN, se utilizaron datos sintéticos generados con base en la topología del área de estudio y modelos teóricos establecidos. Esto permite tener un control preciso sobre los parámetros de entrada y los mecanismos focales conocidos, lo que facilita el proceso de aprendizaje del modelo.

Palabras clave: *mecanismo focal, sismología, evento sísmico, Valle Medio del Magdalena, diseño, inteligencia artificial, red neuronal, convolucional, sintonización, inferencia, estimación, métricas.*

1. INTRODUCCIÓN

La tierra es un sistema complejo y dinámico que se encuentra en constante evolución [1]. Los terremotos son una de las manifestaciones más violentas de su actividad, lo que los convierte en un evento catastrófico capaz de causar daños materiales y pérdida de vidas [2]. Por esta razón, la mitigación de geoamenazas es una prioridad en la investigación en sismología. En este campo, se busca la estimación precisa de los parámetros sísmicos para lograr una mejor comprensión del evento sísmico [3].

La pronta estimación de los parámetros sísmicos, en caso

de un terremoto, es fundamental para comprender la dinámica del evento, evaluar el riesgo de daños a las infraestructuras y a las comunidades cercanas. Además, permite proporcionar una alerta temprana efectiva para salvar vidas y reducir el impacto económico.

Colombia está geográficamente ubicada en el punto de convergencia de las placas tectónicas de Nazca, Sudamérica y Caribe. Debido a la compleja dinámica tectónica de la zona, se originan sismos de diferentes características que pueden tener un alto potencial destructivo [4]. El Valle Medio del Magdalena se encuentra dentro de esta zona, específicamente, al occidente del nido sísmico de Bucaramanga. Este último se define como un volumen de actividad sísmica intensa, persistente en el tiempo y aislado de la actividad de sus alrededores [4]. Debido a su ubicación, el Valle Medio del Magdalena es un lugar idóneo para llevar a cabo estudios y monitoreo constante de la sismicidad en el país.

Actualmente, existen diversas organizaciones dedicadas al desarrollo de métodos para la estimación de los parámetros sísmicos, como el ISNet [5]. La red ISNet está asociada a uno de los departamentos de investigación en el área de sismología de la NIED (National Research Institute for Earth Science and Disaster Resilience) o Instituto Nacional de Investigación de Ciencias de la Tierra y Resiliencia ante Desastres. Actualmente, la NIED se desarrolla como agencia de Investigación ofreciendo análisis sobre eventos naturales como terremotos, volcanes, inundaciones, entre otros [5].

La forma de caracterizar eventos sísmicos se conoce como mecanismo focal. Se trata de una representación bidimensional, vista desde planta, en la que se visualiza la suficiente información para conocer la forma en la que interactúan las placas del subsuelo.

En Colombia, uno de los cinco métodos usados por el Servicio Geológico Colombiano para encontrar el mecanismo focal [6], es el método de las polaridades del primer movimiento, el cual se basa en la polaridad de las ondas P. Si el primer movimiento es hacia arriba es una compresión, al contrario, si es hacia abajo es

una dilatación. Sin embargo, este método requiere la intervención de personal experto que realice el picado de la polaridad. Esto hace que el método se torne lento y tedioso por la cantidad de eventos sísmicos registrados en el Servicio Geológico Colombiano.

La estimación del mecanismo focal suele tener soluciones desde una perspectiva matemática, con aproximaciones de variables físicas y simplificaciones matemáticas que, a pesar de su complejidad y uso respetado a lo largo de los años, pueden presentar poca convergencia con métodos de diferentes autores.

La tecnología de la Inteligencia Artificial (IA) ha revolucionado numerosos campos, demostrando la capacidad de abordar problemas complejos donde no existe método de solución matemático directo o donde los datos de entrada poseen irregularidades como error en toma de datos o ruido. Al exponer a los modelos de IA a una amplia variedad de ejemplos y escenarios, se le entrena para discernir patrones y generalizar a partir de ellos [7].

En el campo de la sismología, la IA ha tenido un impacto positivo, con contribuciones significativas en áreas como la estimación de la magnitud de un terremoto [8], la identificación de fases sísmicas [9], la determinación de tiempos de llegada de la onda P y polaridades de primer movimiento [10], así como la detección y localización de terremotos [11], entre otros. Estos avances se basan en el análisis de datos obtenidos de los sismogramas, los cuales son procesados con el apoyo de modelos de IA para estimar parámetros sísmicos. Estos logros demuestran la notable precisión que se logra al utilizar la IA en comparación con métodos tradicionales, demostrando así ser una herramienta altamente efectiva.

En este proyecto se propone el uso de redes neuronales convolucionales, debido a la progresiva implantación de la IA como solución a problemas de alta complejidad. Este componente ofrece inferencias para la obtención del mecanismo focal a partir de la forma que tenga toda la onda sísmica; a diferencia de algunos métodos convencionales que analizan solo pequeñas partes de la onda. Además, este componente es independiente de la intervención humana para su funcionamiento, lo que reduce posibilidades de error y ralentización de la estimación.

En el presente trabajo se utilizaron datos sintéticos etiquetados con base a los modelos geológicos de la zona del Valle Medio del Magdalena para entrenar una red convolucional. La red se encargará de predecir tres ángulos: Strike, Dip y Rake, expuestos en la Figura 1. Los cuales se utilizarán para construir el mecanismo focal del evento sísmico. De este modo, se obtiene un modelo capaz de estimar el mecanismo focal de un evento sísmico en la zona de estudio con alta precisión y rapidez.

Los resultados obtenidos en este proyecto proporcionan datos confiables que contribuyen al estudio de la dinámica de eventos sísmicos. Estos hallazgos son especialmente valiosos para el monitoreo y la evaluación de los riesgos sísmicos, permitiendo una mejor comprensión de estos fenómenos y su impacto potencial.

2. OBJETIVOS

A. Objetivo General

Diseñar una red neuronal convolucional para estimar los mecanismos focales de sismos sintéticos basados en la topología del Valle Medio del Magdalena.

B. Objetivos Específicos

1. Indagar sobre metodologías para estimar mecanismos focales desde la perspectiva de la inteligencia artificial.

2. Diseñar una topología de red convolucional partiendo de un modelo de red que procese datos sísmicos de similares características.
3. Entrenar y validar la red convolucional diseñada utilizando datos sintéticos etiquetados de la Región del Valle Medio del Magdalena.
4. Sintetizar los hiperparámetros de la topología de red convolucional diseñada que permitan mejorar la estimación del mecanismo focal.
5. Documentar los resultados obtenidos, las métricas de aprendizaje, la topología de datos usados y las características del propio modelo.

3. MARCO REFERENCIAL

Debido a que el proyecto se centra en la solución a problemas relacionados con sismología, es relevante indicar ciertos conceptos básicos.

A. Geometría de la Falla

La geometría de falla se refiere al movimiento relativo de un bloque (bloque techo) respecto a otro (bloque piso) durante un terremoto [12]. La geometría de falla se representa mediante 3 ángulos Strike, Dip y Rake como se representa en la Figura 1.

1. Rumbo o Strike (ϕ): medido desde el norte en sentido horario.
2. Buzamiento de la falla o Dip (δ): medido desde la horizontal hacia abajo en sentido horario respecto al rumbo.
3. Ángulo de deslizamiento o Rake (λ): medido entre una línea horizontal sobre el plano de falla y el vector de deslizamiento.

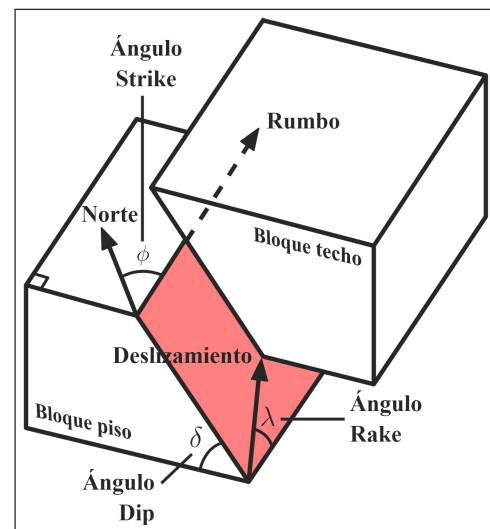


Figura 1. Esquema explicativo del movimiento del bloque techo sobre el bloque piso a través del plano de falla. En el esquema se definen los ángulos que definen la geometría de la falla y el movimiento durante un terremoto. [12]

B. Mecanismos focales

El mecanismo focal es una representación geométrica bidimensional de la geometría de la falla provocada por un evento sísmico, donde se evidencia el movimiento que tuvieron las placas del subsuelo [12].

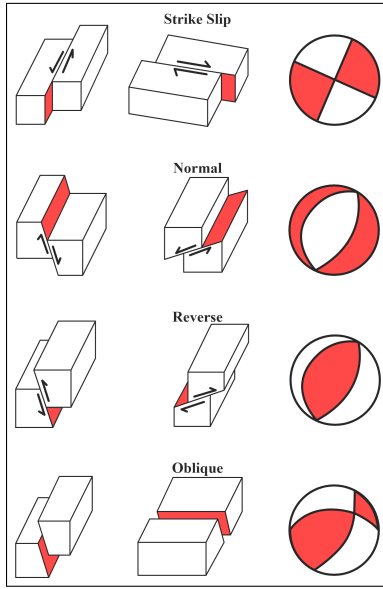


Figura 2. Diferentes mecanismos focales que indican el tipo de movimiento producido por un evento sísmico. [12]

C. Tensor de Momento

Es una representación de la liberación de energía en un evento sísmico donde se proporciona información de la magnitud y dirección. Se calcula realizando mediciones de las ondas sísmicas registradas por los sismógrafos ubicados en distintas estaciones sísmicas [13]. Se puede interpretar como una representación alternativa al mecanismo focal.

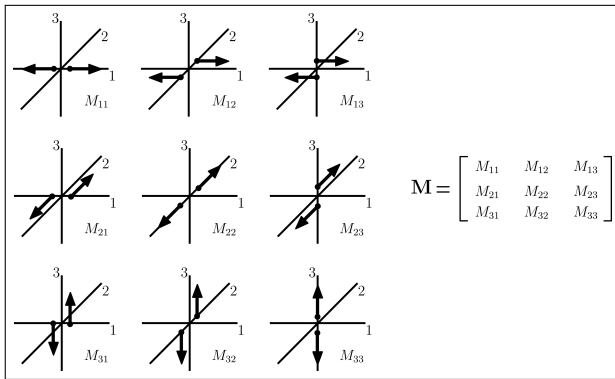


Figura 3. Las nueve componentes resultantes del tensor de momento representadas en el sistema de coordenadas cartesianas. [13]

D. Ángulo de Kagan

Es una medida cuantitativa de la diferencia del ángulo de rotación entre dos mecanismos focales [14].

Teniendo los ángulos Strike, Dip y Rake, se construye la matriz de parámetros de tensor de momento M usando las componentes cartesianas

$$\begin{aligned} M_{xx} &= -M_0[\sin(\delta) \cos(\lambda) \sin(2\phi) + \sin(2\delta) \sin(\lambda) \sin^2(\phi)], \\ M_{xy} &= +M_0[\sin(\delta) \cos(\lambda) \cos(2\phi) + 0.5 \sin(2\delta) \sin(\lambda) \sin(2\phi)], \\ M_{xz} &= -M_0[\cos(\delta) \cos(\lambda) \cos(\phi) + \cos(2\delta) \sin(\lambda) \sin(\phi)], \\ M_{yy} &= +M_0[\sin(\delta) \cos(\lambda) \sin(2\phi) - \sin(2\delta) \sin(\lambda) \cos^2(\phi)], \\ M_{yz} &= -M_0[\cos(\delta) \cos(\lambda) \sin(\phi) - \cos(2\delta) \sin(\lambda) \cos(\phi)], \\ M_{zz} &= +M_0 \sin(2\delta) \sin(\lambda), \end{aligned} \quad (1)$$

$M_{zx} = M_{xz}$, $M_{zy} = M_{yz}$ y $M_{yx} = M_{xy}$ donde M equivale a

$$\begin{bmatrix} M_{zz} & M_{zx} & -M_{zy} \\ M_{xz} & M_{xx} & -M_{xy} \\ -M_{yz} & -M_{yx} & M_{yy} \end{bmatrix}. \quad (2)$$

Para la matriz de tensor de momento de ángulos de etiqueta y estimados se calculan autovectores válidos (V_1 y V_2 respectivamente), posteriormente se calcula el ángulo θ usando la ecuación (3) que corresponde a la diferencia angular entre los dos mecanismos focales. El argumento de arc cos está limitado entre -1 y 1 y tr es la traza del producto punto.

$$\theta = \arccos \left(\frac{tr(V_1 \cdot V_2^T) - 1}{2} \right). \quad (3)$$

E. Métodos de solución para el mecanismo focal propuestos por el Servicio Geológico Colombiano

Se han compilado los métodos disponibles actualmente por parte del Servicio Geológico Colombiano para la estimación del tensor de momento [6]. A continuación se muestran dichos métodos.

1. Fase W: se basa en el análisis de la fase W, la cual se encuentra entre el arribo de la onda P y S [15]. Teniendo la fase W se propone realizar deconvolución, filtrado pasa banda seguido de doble integración, dicho procedimiento está pensado usarse en largos registros de medición.

Posteriormente se realiza un proceso de inversión lineal para la obtención del tensor de momento a partir del tratamiento dado a la fase W [16].

2. ISOLA: el método ISOLA [17], se basa en el análisis de formas de onda completas en base a 6 tensores de momento elementales. Para cada combinación fuente-estación se calculan 6 sismogramas elementales de tres componentes $G(x, y, t)$ convolucionando la función de Green y el tensor de momento elemental. El sismograma debido a un tensor de momento es una combinación lineal de los sismogramas elementales. Donde se realiza un proceso de inversión lineal para la obtención del tensor de momentos.

3. SWIFT: el análisis de las ondas sísmicas en el dominio de la frecuencia ha sido abordado a través del método SWIFT [18], el cual se basa en el análisis de las ondas sísmicas en el dominio de la frecuencia. Para lograr esto, se lleva a cabo un proceso de inversión utilizando la transformada de Fourier. La función de momento se obtiene mediante la aplicación de la transformada inversa de Fourier a las componentes obtenidas en la inversión, para posteriormente pasar por un

proceso de deconvolución desarrollado por el autor. Finalmente, se realiza una búsqueda en cuadrícula para obtener la geometría de la falla, la cual se representa mediante los ángulos de Strike, Dip y Rake.

4. SCMTV: el método SCMTV [19] se fundamenta en base a mediciones de tres componentes alineadas con respecto al arribo de la onda P en una ventana de tiempo predefinida. Estos datos se alinean mediante correlación cruzada con las funciones de Green previamente definidas. A partir de este cálculo, se crea una matriz de inversión y se lleva a cabo el proceso de inversión para resolver el tensor de momento.
5. Polaridades: el método basado en el análisis de polaridades [12], se fundamenta en la identificación de la polaridad del arribo de la onda P en múltiples estaciones de forma simultánea. Dependiendo de la polaridad identificada, se realiza una estimación de la compresión o dilatación experimentada en cada estación. Con esta información, se estima el mecanismo focal que genera dicho comportamiento, tal como se ilustra en la Figura 4.

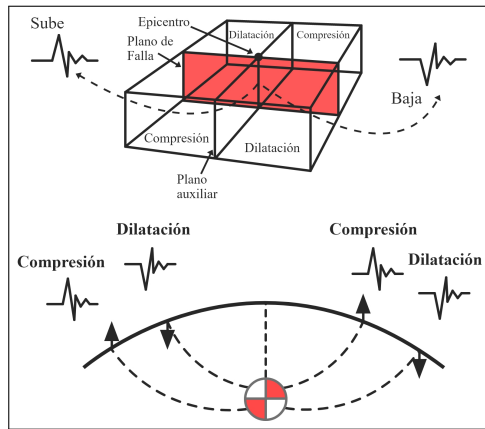


Figura 4. Ejemplo de obtención de un mecanismo focal por el método de primeros arribos. [12]

F. Red Neuronal

Se define como la arquitectura de un modelo computacional donde con unidades conectadas llamadas neuronas, procesan información y aprenden con el ajuste de pesos y conexiones. Esta tecnología puede analizar patrones en datos vistos y realizar procesos de clasificación, estimación o predicción dependiendo su configuración [20]. A continuación, algunas de las clases más comunes de redes neuronales artificiales:

1. Redes Neuronales Feedforward (FNN): son la forma más básica de redes neuronales y consisten en capas de neuronas interconectadas, donde la información fluye en una dirección, hacia adelante.
2. Redes Neuronales Convolucionales (CNN): utilizan capas de convolución para extraer características espaciales y patrones relevantes de las señales. Son ampliamente utilizadas en visión por computadora y tareas relacionadas al procesamiento de señales.
3. Redes Neuronales Recurrentes (RNN): diseñadas para procesar datos secuenciales o con dependencias temporales, como texto, habla o series de tiempo.

4. Redes Neuronales de Memoria de Transformación o Transformer: se destacan en el procesamiento del lenguaje natural y la traducción automática.
5. Redes Neuronales Generativas Adversarias (GANs): consisten en un generador y un discriminador. El generador intenta generar muestras que parecen reales, mientras que el discriminador intenta distinguir entre muestras reales y generadas. Se utilizan para generar contenido sintético, como imágenes y texto.

4. ENTORNO DE TRABAJO

A. Hardware

En el marco de esta investigación, se llevaron a cabo pruebas de entrenamiento utilizando el cluster de la Universidad Industrial de Santander del grupo de Investigación CPS. Este cluster es un recurso indispensable para el procesamiento intensivo requerido en tareas de aprendizaje automático.

El cluster utilizado en este estudio está equipado con dos potentes GPUs (Unidades de Procesamiento Gráfico) para acelerar los cálculos exigentes que implican las múltiples pruebas de la red neuronal. Estas GPUs ofrecen una capacidad de procesamiento paralelo óptima y una memoria de video significativa para manejar grandes conjuntos de datos y modelos complejos. En la Tabla 1 se ilustran las especificaciones.

Tabla 1. Especificaciones de Cluster (Tokio) - CPS

Componente	
GPU 0	Tesla K40c - VRAM 12Gb
GPU 1	Tesla K40c - VRAM 12Gb
RAM	190Gb
Procesador	Intel Xeon E5-2670 v3

B. Software

Conda proporciona un sistema de gestión de paquetes eficiente que permite instalar, actualizar y gestionar las dependencias de manera sencilla, lo que resulta fundamental para la configuración del entorno de entrenamiento de las redes neuronales. La configuración de entorno se ilustra en la Lista 1.

```
$ conda create -n DIREM
$ anaconda python=3.7.7
$ conda activate DIREM
$ conda install ipykernel
$ conda install tensorflow-gpu==2.1.0
$ conda install cudatoolkit=10.1
$ pip install tensorflow==2.1.0
```

Lista 1. Comandos para crear el entorno DIREM e instalar los paquetes

Python 3.7.7 es una versión estable y ampliamente utilizada que ha demostrado ser altamente compatible con numerosas librerías y frameworks de aprendizaje automático, como Tensor Flow y Keras, entre otros.

Después de instalado el framework, se instalaron las librerías compatibles, se especifican las versiones instaladas en la Tabla 2.

Tabla 2. Librerías y versiones Instaladas

Librería	Versión
cupdataoolkit	10.1.243
keras	2.3.1
matplotlib	3.5.3
numpy	1.21.5
pandas	1.3.5
scipy	1.4.1
tensorboard	2.1.1
tensorflow	2.1.0
tensorflow-gpu	2.1.0

Para estas instalaciones se tuvieron en cuenta las características de Software propias del servidor [Tabla 3].

Tabla 3. Software de Cluster

Software	Versión
Debian	4.19.282-1
NVIDIA - Driver Version	418.226.00
CUDA - Version	10.1

5. ÁREA DE ESTUDIO

Colombia, ubicada geográficamente en el punto de convergencia de las placas tectónicas de Nazca, Sudamérica y Caribe, experimenta una compleja dinámica tectónica que genera sismos con diferentes características y un alto potencial destructivo [4]. Dentro de esta zona, se encuentra el Valle Medio del Magdalena, situado al occidente del nido sísmico de Bucaramanga. El nido sísmico de Bucaramanga, conocido por su intensa actividad sísmica, se caracteriza por la ocurrencia de sismos con magnitudes Mw entre 4 y 5, a profundidades que oscilan entre 140 y 200 km [4].

Con una extensión de 32949 km², el Valle Medio del Magdalena se extiende a lo largo del tramo central del valle del río Magdalena, entre las cordilleras Central y Oriental de los Andes colombianos [21]. Esta región, además de su importancia sísmica, es ampliamente explorada y reconocida por albergar numerosos campos de petróleo, convirtiéndola en una zona de gran interés para la explotación de hidrocarburos [21].

Dada su ubicación estratégica y la actividad sísmica asociada, el Valle Medio del Magdalena se presenta como un lugar idóneo para llevar a cabo estudios y un monitoreo constante de la sismicidad en el país.

6. BASE DE DATOS

A. Set de datos

A.1. Modelado de datos sintéticos

En el marco de esta investigación, se emplearon datos sintéticos proporcionados por el grupo de investigación CPS. En esta sección del artículo, se presentan las principales características con las que el grupo generó los datos sintéticos utilizados.

Los datos sintéticos utilizados en este estudio fueron generados teniendo en cuenta el modelo de propagación de ondas elásticas en medios isotrópicos [22] 3D rectangulares.

Para describir la formulación de tensiones [23] y velocidades [24] utilizando el método de diferencias finitas, se resolvió el sistema de ecuaciones en coordenadas cartesianas. La propagación de ondas en la tierra se modeló utilizando la ecuación elástica anisotrópica en tres dimensiones [25].

El grupo de investigación optó por la implementación de la ecuación de onda elástica 3D FDTD (Finite Difference in Time Domain) [25]. De esta manera, el sistema de propagación de ondas se describe mediante un sistema diferencial de primer orden y se discretiza en mallas escalonadas, lo cual simplifica su implementación [26].

$$\begin{aligned}
 \hat{\rho} \frac{\partial V_x}{\partial t} &= \frac{\partial \tau_{xx}}{\partial x} + \frac{\partial \tau_{xy}}{\partial y} + \frac{\partial \tau_{xz}}{\partial z}, \\
 \hat{\rho} \frac{\partial V_y}{\partial t} &= \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \tau_{yy}}{\partial y} + \frac{\partial \tau_{yz}}{\partial z}, \\
 \hat{\rho} \frac{\partial V_z}{\partial t} &= \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \tau_{zz}}{\partial z}, \\
 \frac{\partial \tau_{xx}}{\partial t} &= (\hat{\lambda} + 2\hat{\mu}) \frac{\partial V_x}{\partial x} + \hat{\lambda} \left(\frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} \right) + M_{xx}S(t), \\
 \frac{\partial \tau_{yy}}{\partial t} &= (\hat{\lambda} + 2\hat{\mu}) \frac{\partial V_y}{\partial y} + \hat{\lambda} \left(\frac{\partial V_x}{\partial x} + \frac{\partial V_z}{\partial z} \right) + M_{yy}S(t), \\
 \frac{\partial \tau_{zz}}{\partial t} &= (\hat{\lambda} + 2\hat{\mu}) \frac{\partial V_z}{\partial z} + \hat{\lambda} \left(\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} \right) + M_{zz}S(t), \\
 \frac{\partial \tau_{xy}}{\partial t} &= \hat{\mu} \left(\frac{\partial V_x}{\partial y} + \frac{\partial V_y}{\partial x} \right) + M_{xy}S(t), \\
 \frac{\partial \tau_{xz}}{\partial t} &= \hat{\mu} \left(\frac{\partial V_x}{\partial z} + \frac{\partial V_z}{\partial x} \right) + M_{xz}S(t), \\
 \frac{\partial \tau_{yz}}{\partial t} &= \hat{\mu} \left(\frac{\partial V_y}{\partial z} + \frac{\partial V_z}{\partial y} \right) + M_{yz}S(t),
 \end{aligned} \tag{4}$$

donde $\hat{\lambda}$, $\hat{\mu}$ y $\hat{\rho}$ son parámetros del modelo elástico (parámetros de Lamé y densidad de un sólido isótropo). V_x , V_y , V_z son las velocidades de las partículas en las direcciones x , y , z respectivamente. $S(t)$ es la forma de onda de la ondícula. τ_{xx} , τ_{yy} , τ_{zz} , τ_{xy} , τ_{xz} , τ_{yx} , son los campos de esfuerzos normales y de corte. M_{xx} , M_{yy} , M_{zz} , M_{xy} , M_{xz} , M_{yz} son las componentes del tensor de esfuerzos, asumiendo que: $M_{xy} = M_{yx}$, $M_{xz} = M_{zx}$, $M_{yz} = M_{zy}$ [25].

Se selecciona el enfoque de capa de coincidencia perfecta convolucional (CPML) debido a sus propiedades comprobadas de atenuación y no reflectividad [26]. Se recomienda que la longitud de la zona de atenuación sea al menos de 0.5 a 1 vez la longitud de onda considerada, según se indica en [27].

En cuanto a la fuente sísmica, se optó por implementar un pulso de Ricker como operador para representar la forma de onda de la fuente [28]. El pulso de Ricker tiene una frecuencia central de 0.1 Hz.

A.2. Set de datos

El área de estudio se sitúa en el Valle Medio del Magdalena, abarcando el nido sísmico de Bucaramanga.

La Figura 5 representa el área de estudio que abarca desde -72.4° hasta -74° en longitud y desde 6° hasta 8° en latitud. En cuanto a la profundidad, se considera un rango de 0 a 180 km. Este rango define el volumen de estudio.

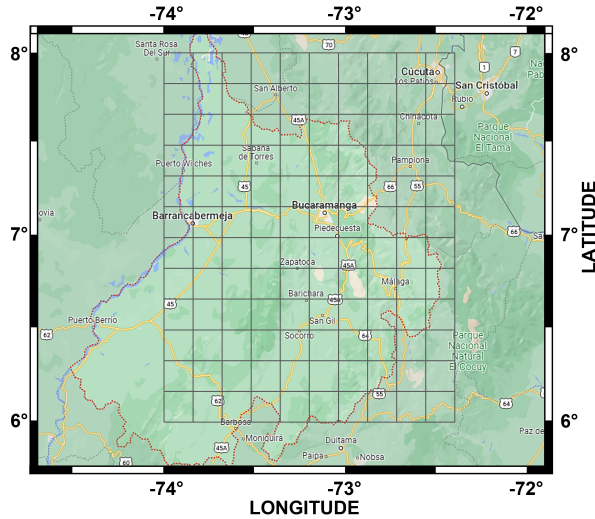


Figura 5. Área de estudio delimitado entre -72.4° a -74° en longitud y desde 6° a 8° en latitud con profundidad en rango de 0 km a 180 km.

Para discretizar el volumen, se han utilizado intervalos de 0.0088° en dirección de longitud, 0.0052° en dirección de latitud y 0.7171 km en dirección de profundidad. De esta manera, el volumen discretizado se muestra en la Tabla 4.

Tabla 4. Discretización de Zona

Dirección	Rango Angular	Rango de muestra
Longitud	-72.4° a -74°	nX [0 - 180]
	paso 0.0088°	paso 2
		90 muestras
Latitud	6° a 8°	nY [0 - 384]
	paso 0.0052°	paso 2
		192 muestras
Profundidad	0 a 180 km	nZ [0 - 250]
	paso 0.7171 km	paso 2
		125 muestras

Dado que la red se ha configurado originalmente para trabajar con 16 estaciones, se ha optado por elegir estas estaciones de manera uniforme, lo que garantiza una cobertura espacial equitativa y representativa de las mediciones. Esto se evidencia en la Figura 6, donde las estaciones están demarcadas de color azul. Esta estrategia busca minimizar posibles sesgos y asegurar resultados confiables en la investigación.

En la región mencionada, se han identificado un total de

75 ubicaciones correspondientes a diferentes eventos sísmicos dentro del nido, los cuales se presentan en la Figura 6 en color rojo. Para estudiar la propagación de ondas en esta área, se ha implementado un modelo de propagación 3D utilizando un pulso de Ricker como fuente en la sección A.1.

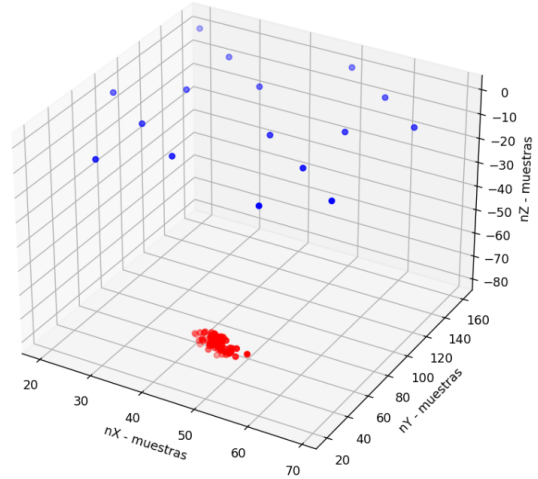


Figura 6. Distribución de la ubicación discretizada de sismos y estaciones disponibles para el estudio.

En el volumen discretizado, se simuló la propagación de ondas sísmicas en las ubicaciones correspondientes a los sismos utilizando un modelo de propagación 3D. Para cada posición del evento sísmico, se generaron formas de onda sintéticas para todas las combinaciones de ángulos mostrados en la Tabla 5 que conforman el mecanismo focal.

Tabla 5. Rangos de orientaciones obtenidos

Dirección	Rango	Paso	Total
Strike	0° a 330°	30°	12 ángulos
Dip	0° a 90°	10°	10 ángulos
Rake	-90° a 90°	20°	10 ángulos

Por lo tanto, se generaron un total de 1200 mecanismos focales para cada posición de los sismos, considerando combinaciones de ángulos. Esto resulta en un total de 90000 (75×1200) datos sintéticos. Estos datos sintéticos corresponden a formas de onda 3-C (componentes de velocidad en las direcciones norte-sur, este-oeste y vertical), registrados por las 16 estaciones durante un intervalo de tiempo de 128 segundos, con una frecuencia de muestreo de 1 Hz.

El set de datos fue normalizado con respecto a la amplitud, quitando así la dependencia de la magnitud del sismo, por lo que la red aprenderá en base a la forma de la onda.

B. Set de etiquetas

Para la implementación de las etiquetas, se decidió tomar los valores de los ángulos que conforman el mecanismo focal (Strike, Dip y Rake) y ajustar cada ángulo en forma de funciones de densidad de probabilidad, en las que la probabilidad máxima de cada distribución corresponde a un componente del mecanismo focal. Estas funciones se ajustaron dentro de los rangos dimensionales con los que puede operar la red, de tal forma que

en post procesamiento se pueda conocer el ángulo estimado, permitiendo la comparativa directa con datos etiquetados.

Ajustar las etiquetas de esta forma ayuda a mejorar la convergencia de la red [29]. Además, la elección adecuada de los parámetros utilizados para construir la distribución gaussiana también es fundamental para lograr un mejor desempeño en el proceso de entrenamiento.

Cada función se ajustó para que pudiera representar de forma simétrica cada uno de los valores alojados en las 3 distribuciones angulares de acuerdo al paso establecido por la base de datos.

La ecuación que describe la función de densidad de probabilidad utilizada para generar la curva en las etiquetas corresponde a

$$f(\hat{x}) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(\hat{x}-\mu)^2}{2\sigma^2}}, \quad (5)$$

donde \hat{x} representa la variable de función, μ representa la media y σ la desviación estándar.

Se calculó el rango significativo de la curva en muestras, teniendo en cuenta que se requiere una sección limitada de esta, considerando un intervalo que compensa adecuadamente la cantidad de muestras requeridas con suavizado en la curva, se llega a que el porcentaje requerido es del 97%. De forma experimental en [29] se concluye que para la arquitectura de red, la desviación estándar óptima es de 5.

En la ecuación (6) se define a w como el número de muestras significativas para la curva y a \hat{x}_1 como la posición donde se obtiene el porcentaje deseado de la altura de la curva.

$$w = 2(\mu - \hat{x}_1). \quad (6)$$

Despejando de la ecuación (5) y reemplazando en la ecuación (6), en la ecuación (7) encontramos un ancho de campana w de 26.48 muestras si se ajusta el 3 % de la altura de la campana (complemento porcentual de 97%). Esto se usará para ajustar los rangos de ángulo al rango finito que admite la red.

$$\hat{x}_1 = \mu - \sqrt{-2\sigma^2 \cdot \ln\left(\frac{3}{100}\right)},$$

$$w = 2(\mu - \hat{x}_1), \quad (7)$$

$$\approx 26.48$$

Dado que cada dirección angular tiene distintos rangos y paso según la Tabla 5, se calcula usando la ecuación (8) los ángulos redondeados adicionales a incluir para completar las campanas.

$$Ajuste = 0.5 \cdot w \cdot \frac{D_{ang}}{128}, \quad (8)$$

donde $Ajuste$ es el ángulo de ajuste y D_{ang} la diferencia entre el ángulo máximo y mínimo del rango de ángulos.

El rango final de ángulos para cada dirección se presenta en la Tabla 6.

Finalmente, en la Tabla 7 se representan las relaciones lineales entre cada dirección angular y el rango de muestras de la etiqueta.

En la Figura 7 se muestra un ejemplo de etiqueta creado.

Tabla 6. Rangos ajustados de orientaciones obtenidos

Dirección	Ajuste	Rango Final
Strike	34°	−34° a 364°
Dip	9°	−9° a 99°
Rake	19°	−109° a 109°

Tabla 7. Ecuación de relacion Angulo - Muestras.

Dirección	Recta
Strike	$\frac{127}{398}\phi + \frac{2159}{199}$
Dip	$\frac{127}{108}\delta + \frac{127}{12}$
Rake	$\frac{127}{218}\lambda + \frac{127}{2}$

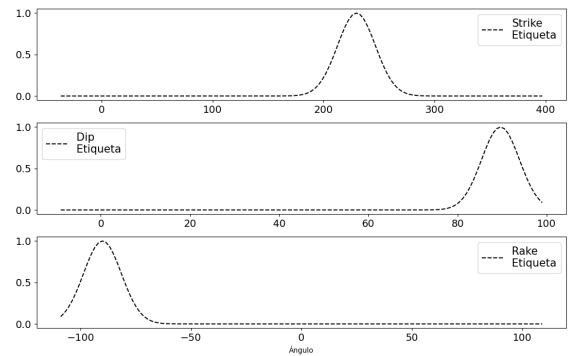


Figura 7. Ejemplo de etiqueta creada para un mecanismo focal con $Strike = 240^\circ$, $Dip = 90^\circ$ y $Rake = -90^\circ$

7. DESCRIPCIÓN DE ARQUITECTURA

A. Estructura de la Red Neuronal Convolutacional

La arquitectura implementada para esta investigación corresponde a una red totalmente convolutacional, conocida como FCN por sus siglas en inglés. Esta arquitectura consta de dos secciones principales: una sección de compresión y una sección de expansión.

A diferencia de las redes convolucionales convencionales que se utilizan comúnmente para clasificar datos en categorías, el modelo FCN utilizado en esta investigación tiene la capacidad de generar una imagen como salida. Esto se logra mediante la combinación de las etapas de compresión y expansión, lo que permite obtener una imagen final que se ajusta a las etiquetas utilizadas en el estudio.

A.1. Diseño de capas

La estructura de las capas convolucionales de la red se basa en un estudio previo [29] que ha demostrado un buen rendimiento acorde a los objetivos de la investigación. La arquitectura consta de 16 capas Conv2D entrenables, 7 capas de MaxPooling, 7 capas de UpSampling, así como 15 capas de LeakyReLU y 15 capas de BatchNormalization. Las cuales se implementan siguiendo el diseño de la Figura 8.

Los datos de entrada tienen una dimensión de 48 x 128 x 1. La dimensión de 48 representa las formas de onda 3-C medidas por las 16 estaciones sísmicas, mientras que 128 es la longitud de la medición. Por otro lado, los datos de salida tienen una

dimensión de $1 \times 128 \times 3$, que corresponde a las tres gaussianas centradas en los ángulos correspondientes al mecanismo focal estimado, mientras que la dimensión de 128 se refiere a la longitud de la predicción.

A.2. Compresión y expansión

Durante el proceso de compresión, los datos de entrada a la red, que tienen un tamaño inicial de $48 \times 128 \times 1$, pasan por capas de MaxPooling, las cuales reducen su tamaño en las dos primeras dimensiones, preservando las características más relevantes. Se utiliza la función de activación LakyReLU y se emplea la técnica de BatchNormalization para normalizar la salida después de que los datos han pasado por el proceso de Conv2D, MaxPooling y LakyReLU.

En este proceso, la capa Conv2D juega un papel fundamental al aumentar el número de filtros o canales en la tercera dimensión. Cada uno de estos filtros tiene un tamaño de 3×3 y desempeñan un papel determinante en el proceso de extracción de características relevantes, siendo indispensables para el proceso de aprendizaje de la red.

El resultado de la fase de compresión es un dato llamado encoder, el cual tiene una dimensión de $1 \times 1 \times 128$. Esta representación reduce la cantidad de información necesaria para describir la entrada original al capturar las características más importantes de la misma.

El encoder desempeña un papel fundamental en el proceso de aprendizaje durante la etapa de expansión, ya que proporciona una base sólida para dicho proceso. Durante esta etapa, los datos atraviesan capas de UpSampling, las cuales aumentan la resolución de las dos primeras dimensiones, generando un incremento en el tamaño de dichas dimensiones en los datos de entrada. Al igual que en el proceso de compresión, se emplean la función de activación LakyReLU y la capa de normalización BatchNormalization.

En este proceso, la capa Conv2D desempeña la función de reducir el número de canales en la tercera dimensión de los datos. Cada filtro utilizado tiene un tamaño de 1×3 . Como resultado de esta operación, se obtiene una salida de dimensiones $1 \times 128 \times 3$, donde 128 es la longitud de la etiqueta y 3 es el número de distribuciones gaussianas creadas para cada ángulo. Durante el entrenamiento, estos datos de salida se ajustaron para obtener tres distribuciones gaussianas. Estas distribuciones representan los ángulos Strike, Dip y Rake, que forman el mecanismo focal del evento sísmico.

8. SINTONIZACIÓN DE PARÁMETROS

Se implementa como función de error cuadrático medio (MSE) como función de pérdida en el entrenamiento, debido a que el aprendizaje de la red se plantea inicialmente como un problema de regresión. Como optimizador, se opta por Adam debido a su capacidad de ajustar automáticamente la tasa de aprendizaje durante el entrenamiento [20].

Dado que el conjunto de datos se produce en base a un nido sísmico que está confinado a un espacio limitado dentro del área de propagación de ondas elásticas, se realizó una revisión espacial de la forma del nido y extrajimos un total de 6 eventos que se categorizaron como representativos de todo el conjunto de datos. Estos sismos se tomaron de tal forma que estuvieran en puntos intermedios y exteriores de toda la forma del nido e incluyendo eventos por cada nivel de profundidad en el espacio discretizado.

Con esto se comprende los puntos de localizaciones principales teniendo en cuenta la densidad del nido. Adicionalmente, se dividió el set resultante en función a las diferentes combinaciones de mecanismos focales, teniendo en cuenta la posición de cada evento sísmico seleccionado. Donde se utilizó un 80 % para entrenamiento y 20 % para validación. Partiremos de establecer el valor por defecto de [29] para poder establecer secuencialmente los mejores valores de Batch y learning rate. En la Figura 9 se representan los sismos empleados sobre la cantidad total.

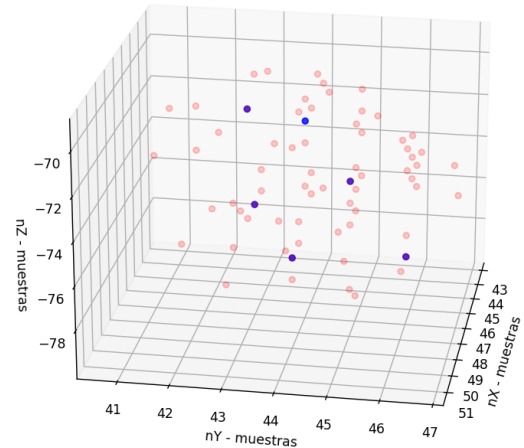


Figura 9. Localización de eventos representativos.

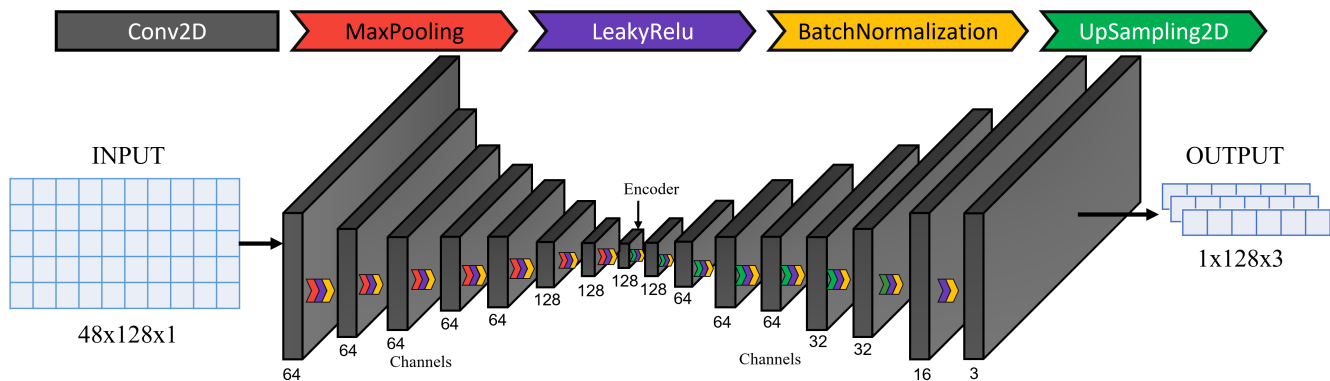


Figura 8. Estructura de la Red Neuronal Convolutional.

A. Ajuste de Batch

Se ejecutó un total de 800 épocas para observar la tendencia de las configuraciones de Batch. Sin embargo, se realizaron entrenamientos de las 6 configuraciones de Batch nuevamente hasta la época 250 por simplicidad dado que desde este punto se sostiene la tendencia de todas las curvas extraídas.

En la Figura 10 se tienen las tendencias en precisión (accuracy) de cada tamaño de Batch. Este comportamiento denota que la red efectivamente tiene un aprendizaje correcto en las primeras épocas, adicionalmente se observa que se estanca alrededor del 65 % para las mejores métricas.

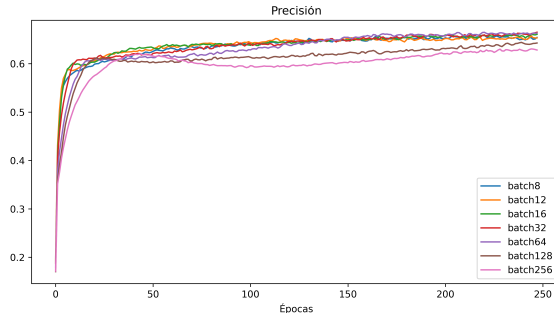


Figura 10. Curvas de precisión para entrenamientos con distintos tamaños de Batch.

En la gráfica de pérdida (loss) mostrada en la Figura 11, se evidencia el decaimiento correcto de todas las mediciones. Las gráficas de precisión y pérdida con respecto al set de validación (validation) podrá encontrarlas en el Anexo A, donde se presenta comportamiento de generalización igualmente correcto para converger con la forma de las métricas de entrenamiento.

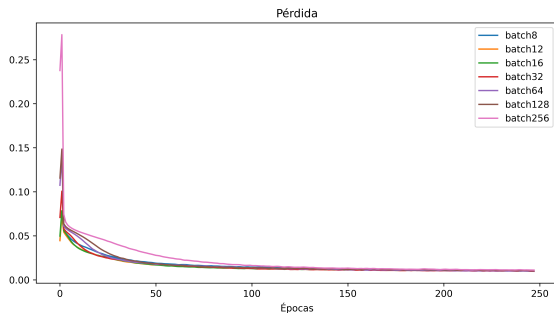


Figura 11. Curvas de pérdida para entrenamientos con distintos tamaños de Batch.

Se evidencia que el tamaño de Batch en los valores más pequeños, se obtienen mejores valores de precisión y pérdida, por el contrario, en la medida que se aumenta el tamaño del batch, los valores tienden a empeorar. Sobre la curva de precisión, tanto en el entrenamiento como en la validación, se destacan tres valores de tamaño de batch: 8, 12 y 16. Estos valores muestran mejores resultados con respecto a los otros cuatro tamaños de lote adicionales que se evaluaron. Adicionalmente se evaluaron los tiempos de ejecución de cada entrenamiento, los cuales se muestran en la Tabla 8 donde se determinó que el mejor tamaño de batch en términos de precisión y tiempo, es de 16.

B. Ajuste de taza de aprendizaje

Ajustado el tamaño de Batch, definimos distintos valores de la taza de aprendizaje (learning rate), los cuales son extraídos

Tabla 8. Tamaños de Batch y tiempos de entrenamiento

Batch	8	12	16	32	64	128	256
Tiempo (min)	96	74	58	33	23	18	16

con base a probar algunos valores por encima y por debajo de los empleados en la red [29]. Estos valores hacen referencia a su estado inicial, el cual se actualiza a medida que avanza el entrenamiento, haciendo uso de los parámetros de momento empleados por defecto. Podemos ver que, tasas de aprendizaje altas pueden tener una convergencia demasiado lenta e incluyendo algunas oscilaciones en mínimos locales, como ocurre con el valor de 0.1, ver Anexo B. Por otra parte, los valores contiguos a 0.001 tiene un buen comportamiento, por lo que podemos tener confiabilidad en este valor, siendo el parámetro por defecto empleado en el optimizador Adam. En la Figura 12 se presentan los valores de tasas de aprendizaje (learning rate) probados.

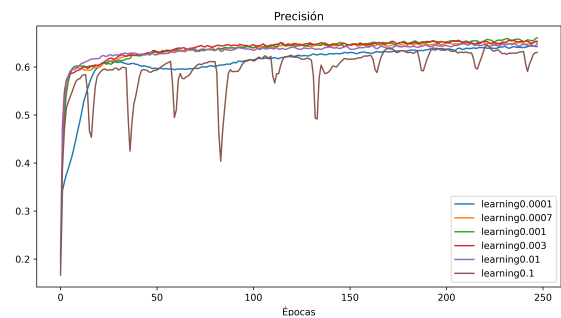


Figura 12. Curvas de precisión para entrenamientos con distintos valores de learning rate.

9. PROCESO DE ENTRENAMIENTO E INFERENCIA

Se dispone de una matriz de datos con dimensiones 90000 x 48 x 128. En esta matriz, el valor 90000 representa la combinación de ángulos (Strike, Dip y Rake) con la cantidad de posiciones de sismos disponibles, lo cual se traduce en 1200 combinaciones de ángulo y 75 posiciones de sismos (1200 x 75 = 90000). El valor 48 corresponde a las diferentes formas de onda 3-C (componentes de ondas sísmicas en tres direcciones) medidas por las 16 estaciones sísmicas (16 x 3 = 48). Por último, el valor 128 indica la longitud de la medición.

En cuanto a las etiquetas, se cuenta con otra matriz de dimensiones 90000 x 128 x 3. En esta matriz, el valor 90000 representa la combinación de ángulos (Strike, Dip y Rake) con las posiciones de los sismos disponibles. El valor 128 se refiere a la longitud de la etiqueta y el valor 3 indica las tres distribuciones gaussianas correspondientes a los ángulos Strike, Dip y Rake que conforman el mecanismo focal.

A. Entrenamiento

Para el proceso de entrenamiento y validación, se han seleccionado 67 sismos de los 75 disponibles, mostrados en la Figura 13. Esto implica que se dispone de un total de 67 x 1200 = 80400 datos para el entrenamiento y la validación. A partir de estos datos, se han destinado 74772 datos para el entrenamiento y 5628 datos para la validación.

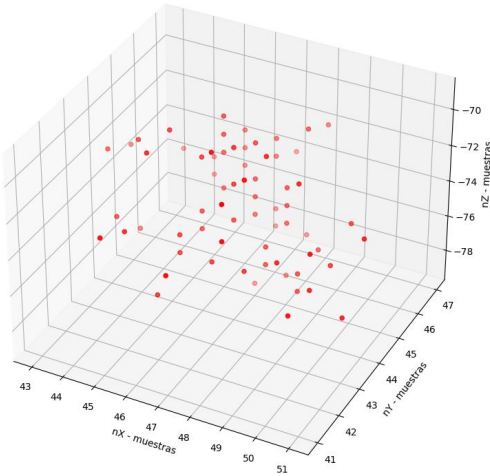


Figura 13. Ubicación de eventos usados en el set de entrenamiento y validación.

Considerando los 90000 datos totales disponibles, se puede calcular el porcentaje correspondiente para el entrenamiento y la validación. Por lo cual, se ha tomado el 83.08 % de los datos disponibles para el entrenamiento ($74772 / 90000 \times 100 = 83.08\%$) y el 6.25 % de los datos disponibles para la validación ($5628 / 90000 \times 100 = 6.25\%$). El restante 10.66 % es destinado para realizar inferencia en la sección C.

Los datos se han mezclado aleatoriamente con el objetivo de evitar posibles sesgos relacionados con la distribución de los sismos. Esta distribución aleatoria se refleja en la Figura 14, la cual muestra la distribución de los ángulos utilizados en el proceso de entrenamiento y validación.

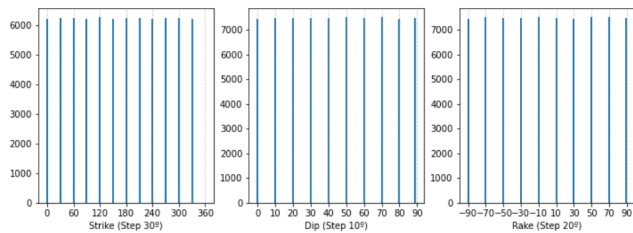


Figura 14. Distribución de ángulos (Strike, Dip y Rake) para el proceso de entrenamiento y validación.

Podemos observar que el modelo examinó los datos de entrenamiento y validación de manera diversa, lo que ayuda a evitar sesgos provenientes del conjunto de datos de entrenamiento original. Esta aleatorización contribuye a garantizar que el modelo se entrene y se evalúe de manera imparcial y equitativa.

B. Métricas de entrenamiento

Para medir la métrica de precisión durante el entrenamiento, se ha implementado la función de pérdida MSE (Mean Squared Error), que compara punto a punto los resultados de la red neuronal con respecto a las etiquetas. A partir de estos cálculos, se generan gráficas de precisión y pérdida generadas en el proceso de entrenamiento, ver Anexo C.

Es importante tener en cuenta que la curva de salida de la red neuronal no es necesariamente idéntica a la curva de entrada, lo cual induce un error al implementar MSE. Sin embargo, se observa que sigue la forma general de la etiqueta, lo cual

indica que la red está aprendiendo a capturar la tendencia y la distribución de los datos.

Con el fin de proporcionar una interpretación más precisa de las predicciones realizadas por la red, se utiliza la métrica de precisión teniendo en cuenta la muestra donde la curva gaussiana alcanza la mayor amplitud. Dado que el etiquetado se discretiza utilizando un paso predefinido, la red intenta realizar predicciones siguiendo ese criterio de paso. Sin embargo, al analizar detenidamente las predicciones de la red, se ha observado que en ocasiones la red falla la predicción por un par de muestras que no están cuantizadas en el conjunto de etiquetas.

Para abordar esta situación, se ha definido un rango de tolerancia al interpretar las predicciones realizadas por la red, considerando que la red puede no ajustarse exactamente a la cuantización predefinida. Al establecer este rango de tolerancia, se busca brindar una mayor flexibilidad en la interpretación de las predicciones y permitir que se considere la precisión en un contexto más amplio.

Para calcular este rango de tolerancia, se hace uso de las ecuaciones que permiten convertir ángulo en muestras, expuestas en la Tabla 7. Estas ecuaciones permiten determinar el paso en muestras para las tres distribuciones gaussianas. Teniendo este paso lo dividimos entre dos para así obtener la máxima tolerancia permitida y aproximamos mediante truncamiento de las cifras decimales, con el fin de obtener una tolerancia máxima adecuada para las mediciones. Dichos resultados se evidencian en la Tabla 9.

Tabla 9. Valor de tolerancias máximas por dirección

Dirección	Paso en ángulo	Paso en muestras (aproximado)	Tolerancia máxima en muestras
Strike	30°	9	4
Dip	10°	11	5
Rake	20°	11	5

Teniendo en cuenta la métrica propuesta, se generan gráficas de precisión para cada dirección de ángulo, Strike, Dip y Rake. Estas gráficas muestran la precisión obtenida en cada dirección de forma individual, ver Anexo D. Promediando las tres precisiones obtenidas, se calcula la precisión total para los datos de entrenamiento. Este resultado se presenta en la Figura 15, donde se observa una precisión media final en torno al 85 %.

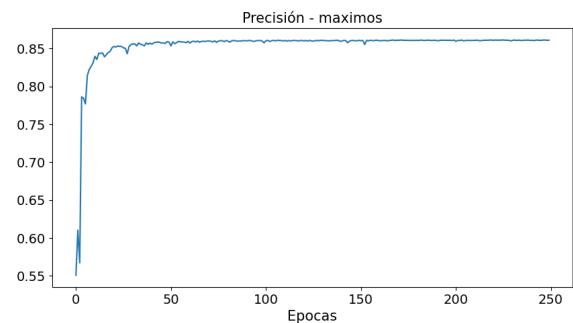


Figura 15. Precisión de entrenamiento en base a métrica con tolerancia.

C. Inferencia

Para el proceso de inferencia, se han seleccionado 8 sismos que la red nunca ha visto antes. Esto implica un total de 9600 datos disponibles, lo que representa aproximadamente el 10.66 % del total de los datos disponibles. Estos sismos adicionales se utilizan para evaluar el rendimiento de la red en situaciones no vistas durante el entrenamiento, para así comprobar su capacidad de generalización más allá de los datos utilizados para entrenarla. Los sismos seleccionados en el área discretizada se presentan en la Figura 16.

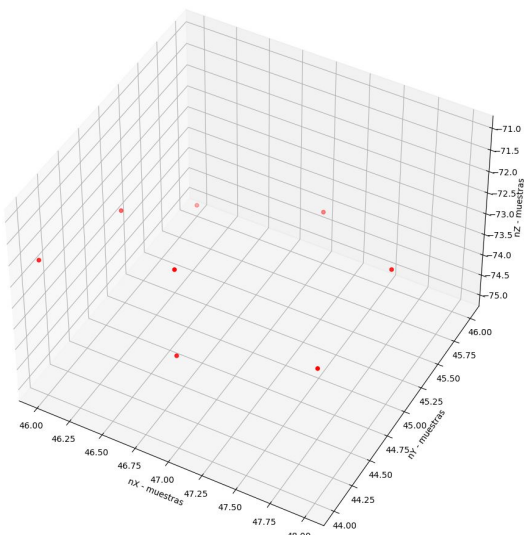


Figura 16. Distribución de eventos usados para el proceso de inferencia.

El modelo seleccionado para realizar el proceso de inferencia se seleccionó obteniendo el valor máximo del promedio de validación, mostrado en el Anexo C. Este enfoque se adoptó debido a que, a pesar de que la curva de promedio de entrenamiento ya se ha estabilizado, aún se observan variaciones en la curva de validación. En base a este análisis se obtuvo que el mejor modelo ocurrió en la época 149. Por lo tanto, se utilizó este modelo para realizar inferencias en adelante.

Los datos usados para el proceso de inferencia siguen la distribución de ángulos mostrada en la Figura 17.

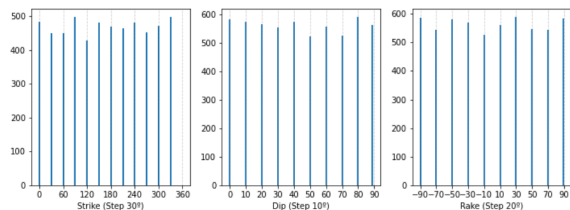


Figura 17. Distribución de ángulos (Strike, Dip y Rake) para el proceso de inferencia.

D. Métricas de inferencia

Durante el proceso de inferencia, se implementa el análisis de histogramas como una métrica para evaluar el rendimiento del modelo. Calculando el porcentaje de error en función de los ángulos Strike, Dip y Rake que conforman el mecanismo focal.

Para este caso, el cálculo del error se basa en la comparación del valor de la amplitud máxima de la curva gaussiana en las muestras predichas con respecto al mismo valor máximo encontrado en las etiquetas correspondientes. Para con ello generar el histograma correspondiente a cada ángulo Strike, Dip y Rake, lo cual se presentan en la Figura 18, Figura 19 y Figura 20, respectivamente.

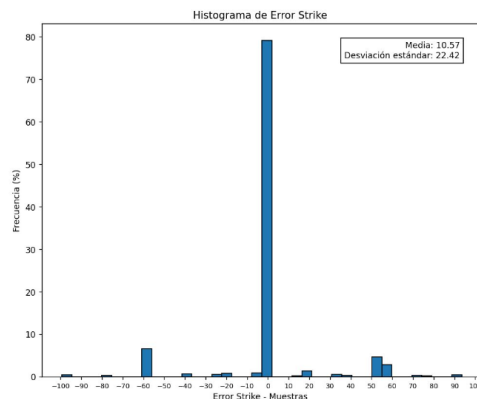


Figura 18. Histograma, media y desviación estándar con relación al ángulo de Strike.

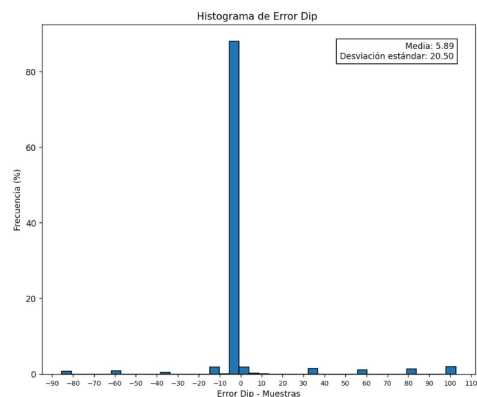


Figura 19. Histograma, media y desviación estándar con relación al ángulo de Dip.

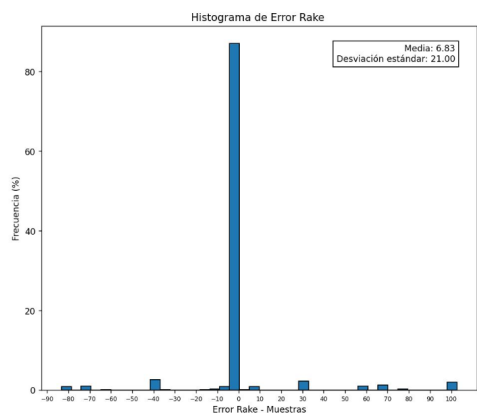


Figura 20. Histograma, media y desviación estándar con relación al ángulo Rake.

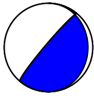

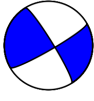
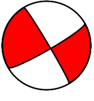


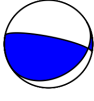
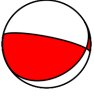
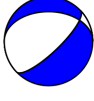
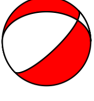
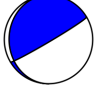
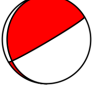
Se calcula el porcentaje de aciertos en función del criterio de tolerancia máxima establecido en la sección B, utilizando los histogramas de error por ángulo. Esto nos permite evaluar la precisión de la inferencia en cada uno de los parámetros del mecanismo focal. Obteniendo así un porcentaje de asertividad de 80.93 % para Strike, 91.98 % para Dip y 88.98 % para Rake.

Con respecto al cálculo de media y varianza, este se calcula en base al valor absoluto del error entre predicción y etiqueta al igual que la desviación estándar.

Para asociar el comportamiento de precisión de la red en sus 3 ángulos de estimación a una estimación correcta o incorrecta, empleamos la comparación entre el mecanismo focal completo estimado y etiquetado. Para lograr esto se emplea el ángulo de Kagan [14] que nos permite estimar cuántos grados diverge un mecanismo focal de otro. En la Tabla 10 se muestra el cálculo del ángulo de Kagan para algunos mecanismos focales estimados por la red.

En el Anexo E se muestran las estimaciones expuestas en la Tabla 10 en forma de curvas de distribución gaussiana para cada combinación de ángulo Strike, Dip y Rake que conforman el mecanismo focal del evento sísmico.

Tabla 10. Algunas predicciones realizadas por la red , comparando la orientación entre balones de playa.

Estimación	Balón Real	Balón Estimado	Ángulo de Kagan
1			4.25°
2			5.14°
3			6.15°
4			0°
5			0°
6			0°

Partiendo de tener todo el set de Test estimado, usamos el recurso de [30] para estimar el ángulo de Kagan entre cada uno de los mecanismos focales, evidenciando los resultados en la Figura 21. Donde adicionalmente se extrajo el porcentaje de estimaciones que tuvieron buena similitud considerando una diferencia angular entre mecanismos menor a 30° según el criterio de Kagan.

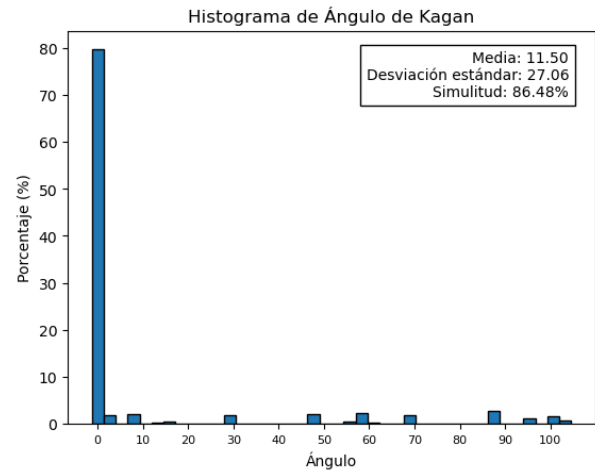


Figura 21. Histograma , media y similitud con relación al ángulo de Kagan.

De acuerdo con el valor de similitud, un 86.48 % de los datos a los cuales la red hizo inferencia, el ángulo calculado entre los balones de playa de etiqueta y predicción, se encuentran por debajo de 30°.

10. DISCUSIÓN

En el marco de la investigación, se demuestra que la implementación del diseño propuesto de la red convolucional, junto con una afinación adecuada de los parámetros, permite lograr una estimación precisa de los ángulos asociados al mecanismo focal de un evento sísmico. Esto se logra al considerar la variación en la forma de onda sísmica , utilizando mediciones de 3 componentes de velocidad, obtenidas de estaciones ubicadas de manera regular en la zona de estudio.

Para el entrenamiento de la red, se hace uso de un conjunto de datos que consiste en formas de onda sintéticas. Como parte del preprocesamiento de estos datos, se realiza únicamente la normalización de su amplitud. Esto permite que la red se enfoque exclusivamente en realizar predicciones basadas en las características de la forma de onda.

Con respecto al etiquetado, se presta especial atención seleccionando una desviación estándar adecuada y configurando la forma de la curva para que fuese suave en todas las medias en las que se centra la curva gaussiana, como se muestra en la Figura 7.

La sintonización de los parámetros se llevó a cabo mediante diversas pruebas, donde se variaron los parámetros para determinar el mejor tamaño de lote (Batch) y tasa de aprendizaje (learning rate), mostrados en la Figura 10 y Figura 12 respectivamente. Obteniendo así, un tamaño de lote de 16 y una tasa de aprendizaje de 0.001 .A partir de estos valores, se realizó el proceso de entrenamiento, el cual arroja una precisión de alrededor del 85 %, evidenciada en la Figura 15. Dicha precisión compara los datos y etiquetas, en base a la posición de los máximos valores de amplitud. Esto se debe a que al calcular la precisión utilizando la función pérdida MSE, se estaría calculando el error como un problema de regresión, lo cual no es lo más adecuado para evaluar la precisión de la red. Incluso si la media de la gaussiana predicha por la red es correcta, siempre habrá un porcentaje de error entre los datos y las etiquetas al calcular punto por punto. Por lo tanto se evaluó la precisión en función de la

posición de los máximos valores tanto de la curva predicha por la red como de la etiqueta.

La red fue puesta a prueba en base a 8 posiciones de sismos nunca vistos, en cada posición se simulaban todas las 1200 posibles combinaciones de los ángulos que conforman el mecanismo focal para posteriormente realizar inferencia sobre los datos mencionados.

Se utilizan histogramas para cuantificar el error entre estimación y predicción en los ángulos de Strike, Dip y Rake, mostrados en la Figura 18, Figura 19 y Figura 20, respectivamente. En todos los casos, se observa que la mayoría de los errores son cercanos a cero, lo cual indica altos niveles de presión para la estimación correspondiente a cada ángulo.

En la Figura 18, Figura 19 y Figura 20 se muestra la media de error para cada ángulo: 10.57 para el Strike, 5.89 para el Dip y 6.83 para el Rake. La media de error para el ángulo de Strike es ligeramente mayor que el paso en muestras, que es de 9 según la Tabla 9. Sin embargo, en el caso del Dip y el Rake, este valor es menor que el paso en muestras, que es de 11 para cada ángulo. Esta diferencia se refleja en el cálculo de las precisiones considerando la tolerancia, donde se obtiene una precisión del 80.93 % para Strike, 91.98 % para Dip y 88.98 % para Rake. Cabe mencionar que el ángulo de Strike presenta la precisión más baja y una media de error más alta en comparación con los otros ángulos.

Esto podría explicarse por el hecho de que el ángulo Strike se discretiza en función a 12 ángulos, mientras que los ángulos Dip y Rake se discretizan en función a 10, según lo mostrado por la Tabla 5. La mayor cantidad de ángulos a representar por parte de Strike, puede afectar la precisión de la estimación. A pesar de esto, en general se observa una alta precisión en la estimación de los ángulos.

Con respecto a la desviación estándar, en los tres casos (Strike, Dip y Rake) muestra una distribución amplia de errores con respecto a la media de error. Esto se refleja en las desviaciones estándar de 22.42 para Strike, 20.50 para Dip y 21.00 para Rake, mostrados en la Figura 18, Figura 19 y Figura 20, respectivamente. Estas desviaciones están directamente relacionadas con las predicciones erróneas generadas por la red.

Para cuantificar el error asociado a la medición basada en los tres ángulos que conforman el mecanismo focal de un evento sísmico, se utilizó el criterio de ángulo de Kagan [14]. Se implementaron histogramas para visualizar la distribución de los ángulos de Kagan, mostrados en la Figura 21, donde se reveló que la mayor concentración de estos ángulos corresponde a valores menores a 30°, lo cual se evidencia en un 86.48 % de similitud. Estos valores se consideran como una buena estimación del mecanismo focal.

En cuanto a las métricas estadísticas mostradas en la Figura 21, se observa que la media de los ángulos de Kagan es de 11.50°, lo cual es menor al criterio de 30°. Esto indica que, en promedio, los errores se mantienen dentro del rango aceptable. Por otro lado, la desviación estándar es de 27.06°, lo que sugiere que existe cierta variabilidad en los errores. Sin embargo, al considerar el criterio de Kagan, esta desviación estándar indica que gran parte de los errores se encuentran dentro del rango aceptable.

El resultado de todo el proceso realizado, es la obtención de un modelo capaz de interpretar la forma de onda y realizar inferencias precisas sobre el mecanismo focal de un evento sísmico. Las métricas de asertividad de la estimación demuestran que en la mayoría de los casos las predicciones son correctas. Estos resultados son alentadores y sientan las bases para futuras

investigaciones.

En futuras investigaciones se puede extrapolar la aplicación de la red a otras zonas con características distintas en sus datos, como frecuencias y ubicación de eventos sísmicos. Además, dado que el etiquetado se realizó utilizando distribuciones gaussianas, en futuras investigaciones se podría explorar la aplicación de paradigmas estadísticos basados en estas distribuciones.

11. CONCLUSIONES

El modelo obtenido, se presenta como un modelo basado en IA, dirigido a la estimación de mecanismos focales, basados en la caracterización de formas de ondas sísmicas. Teniendo como una ventaja destacable, el no requerir de un preprocesamiento riguroso en las mediciones de entrada para su implementación.

La red se modeló con base a datos sintéticos basados en la topología del Valle Medio del Magdalena, una región conocida por su constante actividad sísmica. Estos datos sintéticos proporcionan una representación adecuada de las formas de ondas sísmicas dentro de la zona de estudio. Además, se diseñaron las etiquetas correspondientes con el objetivo de optimizar el proceso de aprendizaje.

En particular se elige una arquitectura basada en CNN óptima para la extracción de características sísmicas. Posteriormente, se llevó a cabo el proceso de sintonización de parámetros de la red diseñada, lo cual permitió que la red adaptara de manera óptima sus pesos.

Durante el proceso de entrenamiento, se emplearon diversas métricas específicamente diseñadas para evaluar el rendimiento del modelo en la estimación del mecanismo focal. El modelo logró capturar de manera efectiva los patrones y las relaciones entre las formas de onda sísmicas y los mecanismos focales asociados. Además, durante el proceso de inferencia, se presentaron sismos nunca antes registrados por la red. Se analizaron las métricas de rendimiento implementadas y se observó que el modelo proporcionó resultados acertados en la estimación del mecanismo focal. Esto demuestra que el modelo puede aplicar los patrones y conocimientos aprendidos durante el entrenamiento a situaciones nuevas y desconocidas. Por lo que se puede concluir que el modelo cuenta con la capacidad para generalizar y discernir patrones a partir de la forma de la onda sísmica, incluso en casos no vistos anteriormente.

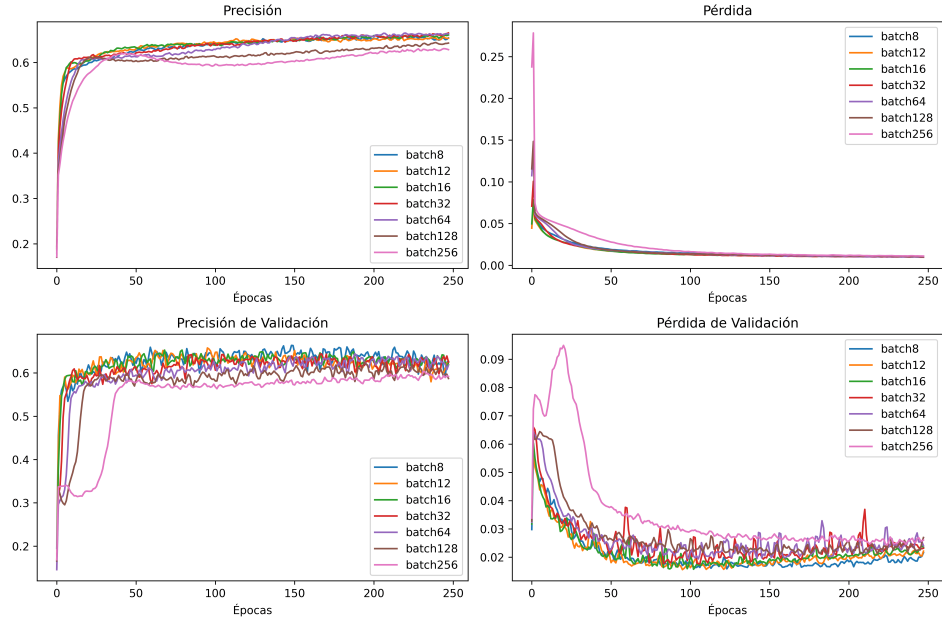
A nivel general, gracias a la investigación llevada a cabo y al diseño propuesto en diversas etapas, se logró desarrollar un modelo altamente preciso. Haciendo hincapié en el preprocesamiento de los datos, prestando especial atención a las etiquetas. Además, se diseñó un modelo de red óptimo para la estimación de mecanismos focales, asegurando una adecuada sintonización de sus parámetros. También se aplicaron métricas de medición existentes y se diseñaron nuevas métricas específicas para evaluar el rendimiento del modelo. Como resultado de este exhaustivo proceso, se obtuvo un modelo capaz de realizar estimaciones con un alto nivel de precisión.

Con base en los resultados obtenidos en este estudio, se brinda una nueva alternativa para la estimación de mecanismos focales mediante el uso de inteligencia artificial, lo que contribuye a la implementación de esta innovadora herramienta en el campo de la sismología. Los resultados obtenidos en este estudio sientan las bases para futuras investigaciones que enriquecerán el análisis de la dinámica de eventos sísmicos, desempeñando un papel fundamental en la evaluación de los riesgos sísmicos y en la búsqueda de una comprensión más profunda de estos fenómenos.

ANEXOS

A. Anexo A

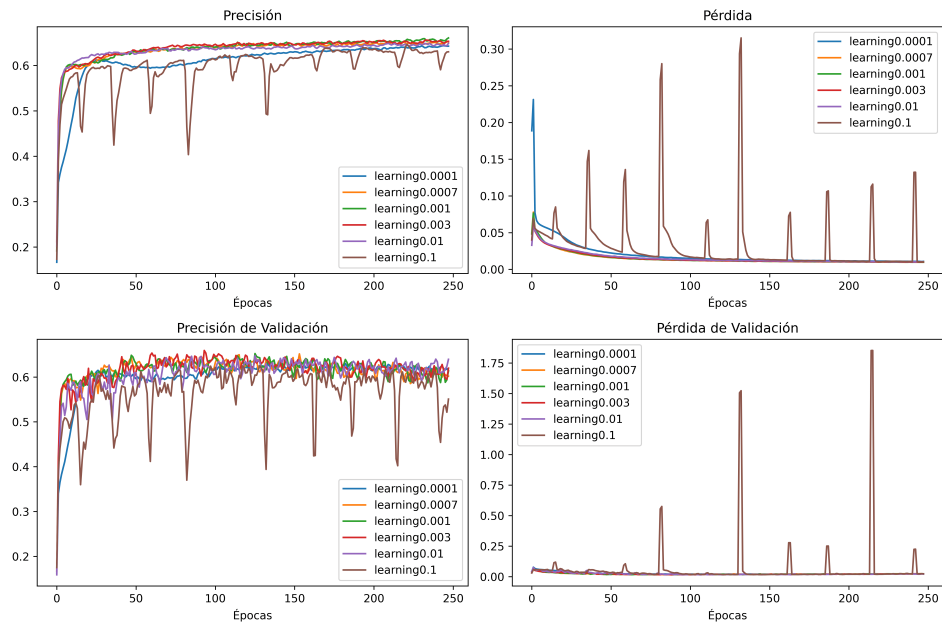
Entrenamientos con ajuste de batch.



Entrenamientos con set de 6 sismos representativos, 250 épocas a un learning rate de 0.001.

B. Anexo B

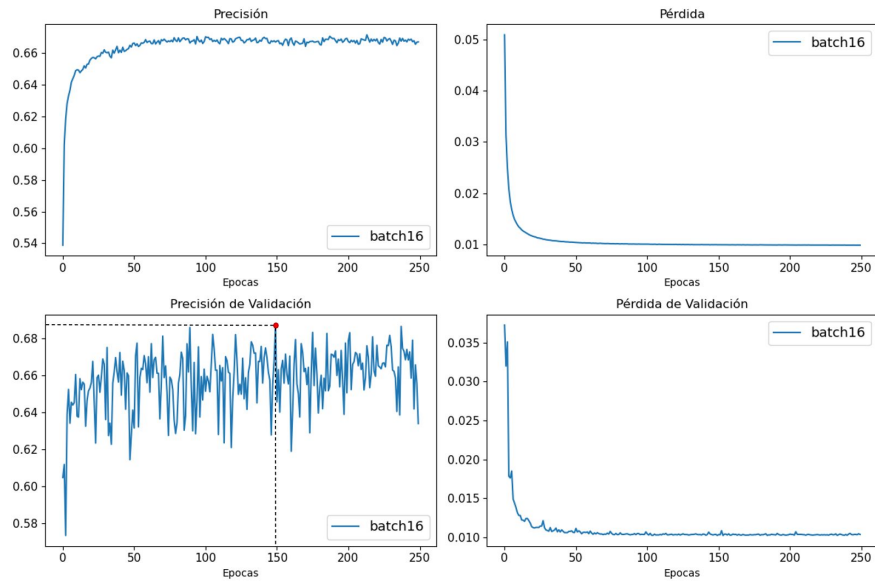
Entrenamientos con ajuste de learning rate.



Entrenamientos con set de 6 sismos representativos, 250 épocas y un tamaño de batch de 16.

C. Anexo C

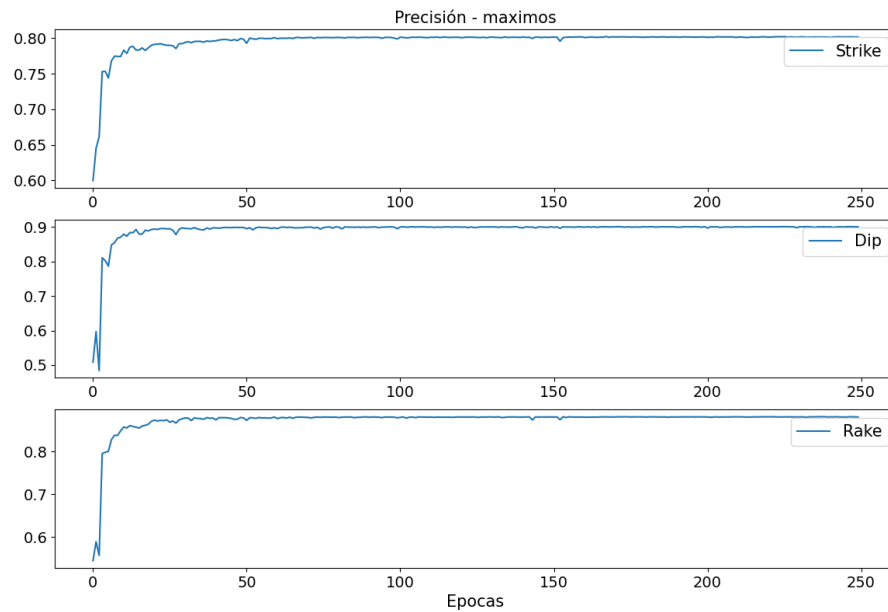
Métricas de entrenamiento a partir de MSE(Mean Squared Error).



Entrenamiento realizado con el set de datos de entrenamiento, 250 épocas y parámetros sintonizados. En la gráfica de precisión de validación, se destaca un punto rojo, el cual representa la época en la que se seleccionó el modelo utilizado en el artículo.

D. Anexo D

Curvas de precisión con tolerancia, correspondiente a los ángulos Strike, Dip y Rake.

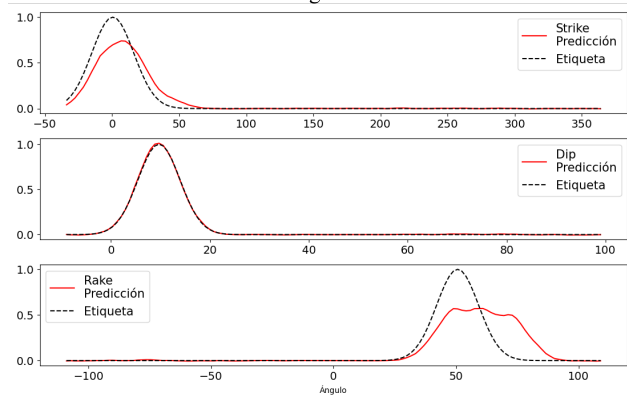


Curvas de precisión con tolerancia para cada ángulo que conforma el mecanismo focal (Strike, Dip y Rake) , 250 épocas.

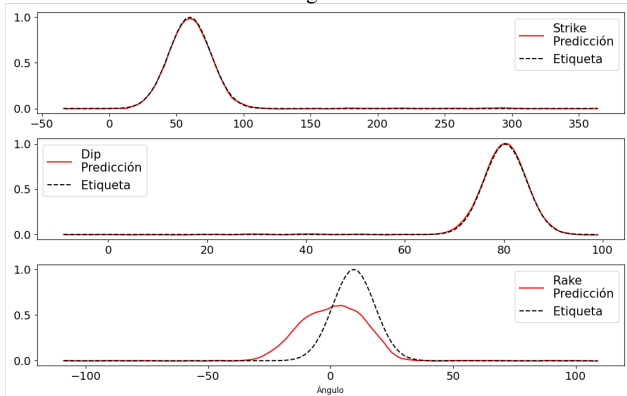
E. Anexo E

Comparativa de estimaciones de la red en base al set destinado para inferencias. Se muestran 3 estimaciones con un porcentaje de error y 3 estimaciones realizadas de manera precisa.

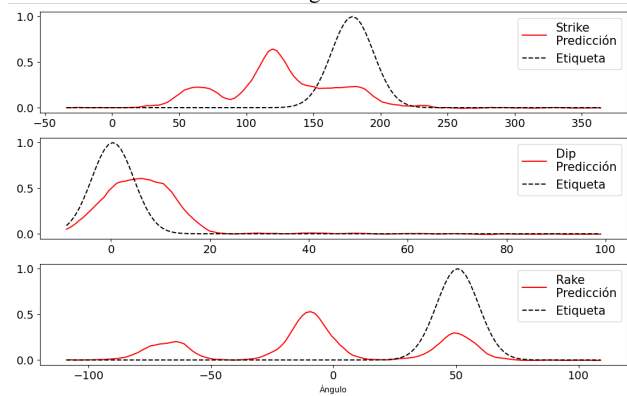
Estimación 1 Diferencia angular de 4.2543°



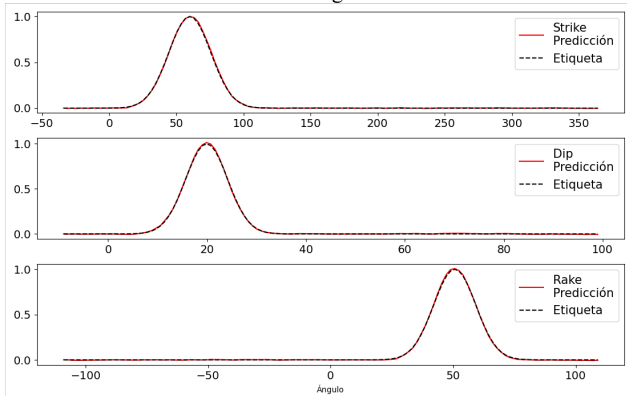
Estimación 2 Diferencia angular de 5.1496°



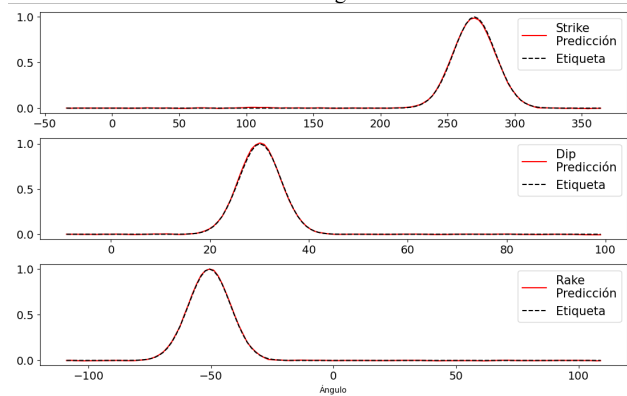
Estimación 3 Diferencia angular de 6.1597°



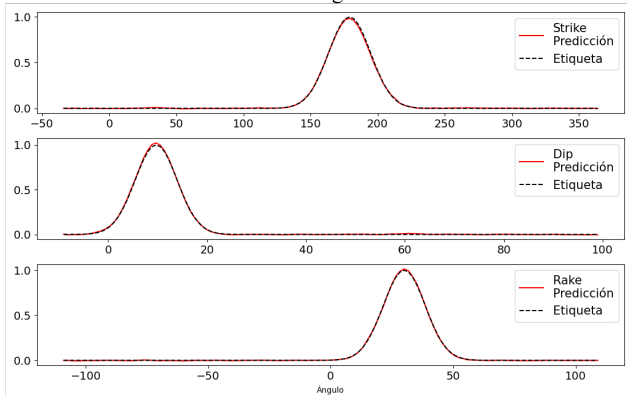
Estimación 4 Diferencia angular de 0°



Estimación 5 Diferencia angular de 0°



Estimación 6 Diferencia angular de 0°



Comparativa de estimaciones de la red en base al set destinado para inferencias.

REFERENCIAS

1. K. Condie, *Earth as an evolving planetary system*, 1st ed. Elsevier Science & Technology, 2005. [Online]. Available: https://uis.primo.exlibrisgroup.com/permalink/57UIDS_INST/ds6hgo/alma991001634694707671
2. G. Salazar, "Terremotos y salud: lecciones y recomendaciones," *Salud Publica Mex*, vol. 60, 2018, doi: <https://doi.org/10.21149/9445>.
3. S. Mazzotti, C. Aubagnac, L. Bollinger, et al., "FMHex20: An earthquake focal mechanism database for seismotectonic analyses in metropolitan france and bordering regions," *Earth Sciences Bulletin*, vol. 192, 2021, doi: <https://doi.org/10.1051/bsgf/2020049>.
4. M. Arcila, J. García, J. Montejó et al., *Modelo nacional de amenaza sísmica para Colombia*, 1st ed. Libros del Servicio Geológico Colombiano, 2020, vol. 43. [Online]. Available: <https://libros.sgc.gov.co/index.php/editorial/catalog/book/38>
5. International Seismic Network of NIED, "About ISNet," https://www.isn.bosai.go.jp/About_ISN.html.
6. Servicio Geológico Colombiano, "Catálogo Mecanismo Focal y Tensor Momento," http://bdrsnc.sgc.gov.co/sismologia1/sismologia/focal_seiscomp_3/index.html#advanced.
7. K. Murphy, *Machine learning a probabilistic perspective*. Cambridge, MA: MIT Press, 2012. [Online]. Available: https://uis.primo.exlibrisgroup.com/permalink/57UIDS_INST/ds6hgo/alma991001686464407671
8. S. Mousavi and G. Beroza, "A Machine-Learning Approach for Earthquake Magnitude Estimation," *Geophysical Research Letters*, 2020, doi: <https://doi.org/10.1029/2019GL085976>.
9. W. Zhu and G. Beroza, "PhaseNet: a deep-neural-network-based seismic arrival-time picking method," *Geophysical Journal International*, vol. 216, 2018, doi: <https://doi.org/10.1093/gji/ggy423>.
10. Z. Ross, M. Meier, and H. Egill, "PWave Arrival Picking and First-Motion Polarity Determination With Deep Learning," *Journal of Geophysical Research*, vol. 123, 2018, doi: <https://doi.org/10.1029/2017JB015251>.
11. T. Perol, M. Gharbi, and M. Denolle, "Convolutional neural network for earthquake detection and location," *Science Advances*, vol. 4, 2018, doi: <https://doi.org/10.1126/sciadv.1700578>.
12. Instituto Geofísico - EPN, "Mecanismos Focales," <https://www.igepn.edu.ec/mecanismos-focales>.
13. S. Cesca, "Inversión del tensor momento sísmico de terremotos superficiales a distancias regionales," 2005. [Online]. Available: <https://eprints.ucm.es/id/eprint/7096/1/T28146.pdf>
14. Y. Kagan, "Simplified algorithms for calculating double-couple rotation," *Geophysical Journal International*, vol. 171, 2007, doi: <https://doi.org/10.1111/j.1365-246X.2007.03538.x>.
15. Centro Sismológico Nacional - Universidad de Chile, "Fase W," <https://www.csn.uchile.cl/fase-w/>.
16. N. Hernández, "Inversión de la fuente sísmica extendida mediante Fase W en el marco del Modelo de Subducción; Una aplicación a grandes terremotos recientes de la zona de subducción chilena," 2019. [Online]. Available: <https://www.dgeo.udec.cl/wp-content/uploads/2019/10/Hernandez-Nicolas-agosto-2019.pdf>
17. J. Zahradník and E. Sokos, "ISOLA Code for Multiple-Point Source Modeling—Review," *Springer International Publishing*, 2018, doi: https://doi-org.bibliotecavirtual.uis.edu.co/10.1007/978-3-319-77359-9_1.
18. M. Nakano, H. Kumagai, and H. Inoue, "Waveform inversion in the frequency domain for the simultaneous determination of earthquake source mechanism and moment function," *Geophysical Journal International*, vol. 173, 2008, doi: <https://doi.org/10.1111/j.1365-246X.2008.03783.x>.
19. AUTOMT/MTV, "Moment tensor inversion documentation," <https://docs.gempa.de/mt/current/apps/scmtv.html>.
20. A. Géron, *Hands-On Machine Learning With Scikit-Learn, Keras, And Tensorflow: Concepts, Tools, And Techniques To Build Intelligent Systems*, 3rd ed., Seapostol, California, Ed., 2022. [Online]. Available: https://uis.primo.exlibrisgroup.com/permalink/57UIDS_INST/ds6hgo/alma991001713201107671
21. Agencia Nacional de Hidrocarburos, "Valle Medio del Magdalena," <https://www.anh.gov.co/es/hidrocarburos/oportunidades-disponibles/procesosde-seleccion/ronda-colombia-2010/tipo-1/>
22. Pedro Pablo Gil-Crespo, Universidad del País Vasco, "Isotropía y anisotropía óptica: indicatrices," <https://www.ehu.es/mineralogiaoptica/isotropia-y-anisotropia-optica-indicatrices.html>.
23. Gobierno de México, "SISMOS: Causas, características e impactos," <https://www.gob.mx/sgm/articulos/sismos-causas-caracteristicas-e-impactos?idiom=es>.
24. Instituto Geográfico Nacional, "Sismología - ¿Qué es la Sismología?" <https://www.ign.es/web/sis-teoria-general>.
25. J. Luna, "Realizar módulo que permita calcular los campos hacia adelante y hacia atrás de la ecuación de onda elástica," *CPS Res. Group*, 2022.
26. C. A. Pérez Solano, "Two-dimensional near-surface seismic imaging with surface waves: alternative methodology for waveform inversion," *ParisTech*, 2013. [Online]. Available: <https://pastel.hal.science/pastel-00932790#>
27. D. Komatitsch and R. Martin, "An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation," *Geophysicists*, vol. 72, 2007, doi: <https://doi.org/10.1190/1.2757586>.
28. M. García, "Cálculo de sismogramas sintéticos mediante el método de número de onda discreto en un medio estratificado (caso SH)," *Universidad Nacional Autónoma de México, Facultad de Ingeniería*, 2009. [Online]. Available: <https://repositorio.unam.mx/contenidos/3531885>
29. W. Kuang, C. Yuan, and J. Zhang, "Real-time determination of earthquake focal mechanism via deep learning," *Nature Communications*, vol. 12, 2021, doi: <https://doi.org/10.1038/s41467-021-21670-x>.
30. M. Hearne et al., "SeismoTectonic Regime Earthquake Calculator (STREC)," <https://github.com/usgs/strec>.