

Detección automática de primeros arribos en trazas sintéticas terrestres del Valle Medio del Magdalena usando redes neuronales artificiales.

Oscar Fabian Reyes Cachopo y Silvia Janeth Rincón Zambrano

Director

Óscar Mauricio Reyes Torres

Doctor en Ingeniería Electrónica

Codirector

Sergio Alberto Abreo Carrillo

Doctor en Ingeniería Electrónica

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingenierías Eléctrica, Electrónica y de Telecomunicaciones

Bucaramanga

2022

Dedicatoria

Este gran logro, en especial y con mucho cariño, lo dedico a mi hermano Victor por su apoyo incondicional en la construcción de mi vida profesional, sus infinitas virtudes y su gran corazón me llevan a admirarlo cada día más.

A mi madre y padre, quienes amo con todo mi ser, infinitas gracias por ser parte de mí, su amor y cariño fueron clave para este resultado. Mamá mil gracias por tu fortaleza, eres y serás mi mayor admiración en la vida, tu esfuerzo es impresionante y tu amor es para mí invaluable.

A mis amigos Juan, Paula, Pablo, Daniela, Sergio, Edward, Karen F, Roman, Karen A, Danilo, Sebastián, Karen G, Felipe y Alan, muchas gracias por su compañía en esta trayectoria.

Muchas gracias por ser parte de mí y ayudarme a culminar esta meta, con todo cariño dedico a todos ustedes este gran logro.

Oscar Fabian Reyes Cachopo

A mis padres Mary y Jose por darme la vida, por su amor, compañía y paciencia incondicional, por inculcar en mí la constancia y perseverancia en cada reto que decido emprender, por todos los sacrificios que han hecho por brindarnos lo mejor a mí y mis hermanos, no me alcanzará la vida para agradecerles; a mis hermanos Leonardo, Kamila y Gerónimo por mantener en mí la fuerza y convicción de seguir luchando por mis sueños, por acompañarme a lo largo de esta historia, sin ustedes nada de esto sería posible.

A Sergio, mi compañero de aventuras, por brindarme su amor y por estar siempre cuando más lo necesitaba, espero que Dios nos permita seguir construyendo un futuro juntos. A mis mejores amigos, Karen, Juan, Laura, Sergio, Gaby, Dalai, Javier, Ingrid, Tatiana, Jose y Valentina, por ser mi compañía a lo largo de este camino, por permitirme ser, crecer y aprender, ustedes son parte de mí, les deseo siempre lo mejor, aún tenemos una promesa por cumplir, Dios les bendiga en todo lo que realicen. A Laura, de las coincidencias más bonitas de la Universidad, quisiera poder explicar en palabras cuán agradecida estoy; a todos ustedes con profundo cariño dedico esto.

Silvia Janeth Rincón Zambrano

Agradecimientos

Agradezco a Dios por la vida y la oportunidad de formarme de manera íntegra para servirle a la sociedad. Agradezco a mi familia por acompañarme con su amor a lo largo de este camino. A todos mis compañeros, docentes y miembros de la Universidad Industrial de Santander por ser parte de este proceso de formación y aprendizaje académico.

Mi más sincero agradecimiento a los docentes Sergio Abreo y Óscar Reyes por su compañía, enseñanza y paciencia durante el desarrollo de este proyecto; gracias por enseñar desde la vocación, por incentivar en los estudiantes el interés por el conocimiento de la ciencia y tecnología, pero sobretodo, gracias por inspirarnos y motivarnos al ver en cada uno de nosotros un colega sin distinción alguna. A Oscar, mi compañero de tesis, por su entrega y arduo trabajo a lo largo de este proyecto, ha sido una experiencia con un sinnúmero de aprendizajes, muchos éxitos en su vida profesional. Al grupo de investigación en Conectividad y Procesamiento de Señales (CPS), a la empresa petrolera colombiana Ecopetrol y al geofísico William Agudelo por ser partícipes en desarrollo y culminación de este proyecto.

Finalmente, a Mauricio, María Fernanda y Adriana por ser mi compañía espiritual a lo largo de mi pregrado, por brindarme herramientas para confiar en mis capacidades y superar las adversidades que se presentaron. A todos ustedes: ¡Mil gracias!.

Silvia Janeth Rincón Zambrano

Tabla de Contenido

Introducción	16
1 Objetivos	20
1.1 Objetivo General	20
1.2 Objetivos Específicos	20
2 Marco de referencia	21
2.1 Contexto geofísico	21
2.1.1 Método sísmico	21
2.1.2 Onda elástica	23
2.1.3 Trazas sísmicas	24
2.1.4 Atributos sísmicos	24
2.1.5 Detección de primeros arribos	26
2.2 Inteligencia artificial	28
2.2.1 Algoritmos implementados para análisis de señales sísmicas	29
2.2.1.1 Redes neuronales	30
2.2.2 Función de Activación.	32
2.3 Entornos de trabajo	32
2.3.1 Hardware	33

DETECCIÓN AUTOMÁTICA DE PRIMEROS ARRIBOS.	7
2.3.2 Anaconda	33
2.3.3 Librerías y paquetes	34
3 Prototipos implementados para selección de la estrategia de detección	36
3.1 Primera estrategia: Detección de primeros arribos basada en múltiples transformadas	41
3.2 Segunda estrategia: Detección de primeros arribos basada en transformadas Wavelets	56
3.3 Selección de Topología	67
4 Red solución para la detección de primeros arribos	72
4.1 Experimentos y comparación de resultados	73
4.1.1 Experimentos red convolucional	74
4.1.2 Variación de parámetros e hiperparámetros de la CNN unida con la DNN	81
4.2 Red solución para predecir el primer arribo	88
5 Conclusiones	103
6 Recomendaciones	107
Referencias Bibliográficas	109
Apéndices	115
A Códigos de programación implementados en estrategias y red solución	115

Lista de Figuras

Figura 1	Método de adquisición en señales sísmicas	22
Figura 2	Atributos sísmicos	26
Figura 3	Trazas sísmicas con su respectivo primer arribo	27
Figura 4	Estructura de una red neuronal	31
Figura 5	Funciones de activación	32
Figura 6	Características de recursos en Hardware	33
Figura 7	Representación gráfica del inventariado antes y después del primer arribo	37
Figura 8	Traza sísmica normalizada con adición de ruido	39
Figura 9	Entropía de una traza sísmica	43
Figura 10	Arquitectura DNN para la red pre y pos primer arribo	43
Figura 11	Primer arribo verdadero y primer arribo falso sobre una traza sísmica	44
Figura 12	Arquitectura DNN para la detección del primer arribo	45
Figura 13	Proceso realizado por las redes DNN pre y pos primer arribo	46
Figura 14	Inventariado pre y pos primer arribo con etiqueta verdadera y falsa	47
Figura 15	Función de pérdidas de entrenamiento y validación en las redes DNN pre y pos primer arribo	49
Figura 16	Distribución de datos de entrenamiento y validación con diferentes niveles de SNR	51

Figura 17	Salida en las redes DNN pre y pos primer arribo	52
Figura 18	Identificación del primer arribo predicho en la salida de las redes pre y pos primer arribo	53
Figura 19	Detección del primer arribo predicho por las redes DNN pre y pos primer arribo	53
Figura 20	Pérdidas de entrenamiento y validación en la red pre y pos primer arribo	55
Figura 21	Arquitectura CNN para la red de detección del primer arribo	57
Figura 22	Entrenamiento red convolucional	59
Figura 23	Respuesta en frecuencia filtro pasa baja	60
Figura 24	Traza sísmica con una de sus componentes Wavelet	63
Figura 25	Conjunto de transformadas Wavelet para una traza sísmica, red CNN	63
Figura 26	Salida de la red CNN	64
Figura 27	Estimación del valor del primer arribo con la red CNN	65
Figura 28	Función de pérdida de entrenamiento y validación de la red CNN	66
Figura 29	Distribución de la predicción de primeros arribos para las redes CNN y DNN	68
Figura 30	Histograma con porcentajes de error en la predicción de primeros arribos para las redes CNN y DNN	69
Figura 31	Arquitectura red solución: unión entre CNN y DNN	78
Figura 32	Funciones de pérdida en entrenamiento y validación de las variaciones de estructura en la unión de CNN con DNN	80
Figura 33	Función de pérdida de la unión de las redes CNN con DNN	81

Figura 34	Histogramas con métricas de error en predicción del FB de la “Serie a” de la	
	Tabla 11	86
Figura 35	Histogramas con métricas de error en predicción de la serie de “Serie b” de la	
	Tabla 11	88
Figura 36	Arquitectura de la red solución	89
Figura 37	Etiqueta de entrenamiento en la red solución	90
Figura 38	Representación gráfica de la salida y predicción de la red solución	91
Figura 39	Estimación del valor del primer arribo con la red solución	93
Figura 40	Función de pérdida de entrenamiento y validación red solución.	94
Figura 41	Histograma del desempeño de la red solución en la predicción de primeros arribos	96
Figura 42	Histograma de distribución de primeros arribos red solución	97
Figura 43	Histograma de distribución de primeros arribos red solución para datos con	
	SNR variable	99
Figura 44	Histograma de distribución para trazas con primeros arribos de 0 a 50 muestras	102
Figura 45	Diagrama de flujo del procesamiento de la detección de primeros arribos	117

Lista de Tablas

Tabla 1	Conjunto de datos sísmicos en formato SEG Y	38
Tabla 2	Definición de parámetros para la arquitectura DNN	49
Tabla 3	Cálculo de métricas para las redes pre y pos FB con datos de entrenamiento y validación	54
Tabla 4	Conjunto de datos sísmicos para red CNN	58
Tabla 5	Definición de parámetros para la arquitectura CNN	60
Tabla 6	Métricas de resultados para red CNN	66
Tabla 7	Diseño de experimentos para la selección de parámetros de la CNN	75
Tabla 8	Errores con datos de entrenamiento y validación, red CNN	76
Tabla 9	Métricas de resultados variando estructuralmente la unión de CNN con DNN	79
Tabla 10	Análisis de sensibilidad en la identificación de parámetros e hiperparámetros	84
Tabla 11	Errores variando parámetros e hiperparámetros de la unión de CNN con DNN	85
Tabla 12	Definición de hiperparámetros para la arquitectura de solución propia	92
Tabla 13	Selección de parámetros para el diseño de la red solución	92
Tabla 14	Definición de hiperparámetros para la arquitectura DNN usada en red solución	92
Tabla 15	Métricas de resultados para red solución con datos de entrenamiento y validación	94
Tabla 16	Métricas de resultados para red solución con datos de validación con variación de SNR	98

Tabla 17	Resultados de la predicción de primeros arribos para trazas de sólo ruido.	101
Tabla 18	Métricas de resultados para red solución con datos de validación	101

Lista de Apéndices

Apéndice A. Códigos de programación implementados en estrategias y red solución

Resumen

Título: Detección automática de primeros arribos en trazas sintéticas terrestres del valle medio del Magdalena usando redes neuronales artificiales.

*

Autores: Oscar Fabian Reyes Cachopo, Silvia Janeth Rincón Zambrano. **

Palabras Clave: Redes Neuronales, Picado de primeros arribos, Geofísica, Trazas Sísmicas, TensorFlow, Keras, Transformada Wavelet.

Descripción: Este trabajo se centra en la detección de primeros arribos de un conjunto de datos sísmicos generados sintéticamente con características del Valle Medio del Magdalena mediante el uso de redes neuronales artificiales. Para ello, se estudian los diferentes conceptos y elementos que conforman el evento sísmico permitiendo la obtención de atributos que entregan información en la detección del primer arribo. Dentro de la adquisición sísmica, la presencia de ruido enmascara el pulso que corresponde al primer arribo, lo que dificulta la tarea de detección y hace necesario que un experto realice esta labor de forma manual. Sin embargo, el gran volumen de datos sísmicos, el tiempo requerido y los diversos criterios de selección basados en la subjetividad del operador, producen picados sesgados que aumentan la posibilidad de errores durante el proceso, haciendo adecuado el uso de herramientas que faciliten la detección de forma automática. La arquitectura de la red solución es obtenida a partir de la unión de dos estrategias encontradas en el estado del arte. La primera plantea una detección a partir de información de la traza cruda y un conjunto de transformaciones mediante el uso de redes densamente conectadas. La segunda consiste en una red convolucional la cual realiza la detección haciendo uso de filtro pasa bajas, la traza cruda y 31 transformaciones Wavelet. Estas estrategias determinan los primeros arribos analizando las trazas sísmicas individualmente. Nuestra propuesta consiste en una red solución conformada por una secuencia de capa convolucional 1D, *maxpooling* 1D, capa *flatten* y dos capas densas. Los parámetros e hiperparámetros de la red son seleccionados mediante múltiples iteraciones y experimentos, llegando así a la identificación del tiempo de llegada de la onda refractada con el uso de transformadas Wavelet, filtro pasa bajas y las trazas sísmicas normalizadas. Finalmente con esta solución se logra alcanzar aciertos con un error máximo de cinco muestras en el 81,08% de 4108 trazas sísmicas con 1251 muestras y niveles de ruido entre el rango de 6 dB a 10 dB. Adicionalmente se evalúa la red variando la relación señal a ruido entre 0, 3, 6, 10 y 20 dB, proporcionando aciertos del 61,85% , 72,03% , 77,67% , 82,78% y 85,10% respectivamente con errores máximos de 5 muestras frente al arribo verdadero.

* Trabajo de Grado: Detección automática de primeros arribos en trazas sintéticas terrestres del valle medio del Magdalena usando redes neuronales artificiales.

** Facultad de Ingeniería Fisicomecánicas. Escuela de Ingenierías Eléctrica, Electrónica y Telecomunicaciones. Director: Óscar Mauricio Reyes Torres. Codirector: Sergio Alberto Abreo Carrillo

Abstract

Title: Automatic detection of first arrivals in synthetic land traces of the middle Magdalena valley using artificial neural networks. *

Authors: Oscar Fabian Reyes Cachopo, Silvia Janeth Rincón Zambrano. **

Keywords: Neural Networks, First Arrival Picks, Geophysics, Seismograms, Seismic Tracing, TensorFlow, Keras, Wavelet transform.

Description: This work focuses on the detection of first arrivals of a synthetically generated seismic data set with characteristics of the Middle Magdalena Valley through the use of artificial neural networks. For this purpose, the different concepts and elements that make up the seismic event are studied, allowing the obtaining of attributes that provide information in the detection of the first arrival. Seismic exploration makes it possible to acquire subsurface information that indicates the behavior of the earth's layers from the propagation of naturally or artificially generated waves. The data received are recorded by seismographs, in the form of signals called seismic traces. As part of the seismic data acquisition process, the first arrival is detected. First arrival detection consists of determining the arrival times of the first refracted wave in each trace. The first arrival is exposed to noise that can be produced by the source, the receiver or the medium where the disturbance propagates. However, the presence of noise as well as other propagating waves masks the pulse corresponding to the first arrival, thus hindering the detection task. Traditionally this process is carried out by an expert who performs this task manually. However, the large volume of seismic data, the time required and the various selection criteria based on operator subjectivity, produce biased picks that increase the possibility of errors during the process. The architecture of the solution network is obtained from the union of two strategies found in the state of the art. The first one proposes a detection based on raw trace information and a set of transformations using densely connected networks. The second consists of a convolutional network which performs detection using low-pass filtering, the raw trace and 31 Wavelet transforms. These strategies determine the first arrivals by analyzing the seismic traces individually. In this work, we propose an automatic first-break picking strategy based on neural networks to detect the first arrival on data with characteristics of the Middle Magdalena Valley in Colombia. We use a neural network that combines a 1D convolutional layer and two dense layers to classify the samples of a preprocessed trace between pre- and post-first arrival, to finally determine the most likely sample corresponding to the first-break in a postprocessing step. Our strategy allows detecting the first arrival in 95.86%, 91.19% and 77.36% of the cases, with a margin of error of 10 samples, when the SNR is 20 dB, 6 dB and 0 dB, respectively, proving to be a suitable strategy for noisy signals.

* Undergrad Work

** Faculty of Engineering Physicomechanics. School of Electrical, Electronic and Telecommunications Engineering. Advisor: Óscar Mauricio Reyes Torres; Co-Advisor: Sergio Alberto Abreo Carrillo

Introducción

La adquisición de información sísmica ocurre a partir de la generación de perturbaciones ocasionadas por una fuente, las cuales se propagan por el medio terrestre y son registradas en superficie a través del uso de receptores, captando frentes de onda que arriban por efectos de refracción y reflexión, estos registros reciben el nombre de trazas sísmicas. El conjunto de trazas organizado por una característica en común recibe el nombre de sismograma, las trazas sísmicas representan una serie de tiempo que registra información sobre características y atributos del subsuelo como la densidad, velocidad, viscosidad, entre otras (Brown, 2011). Durante el procesamiento de datos sísmicos, la detección del primer arribo consiste en determinar el primer tiempo de llegada de la señal en cada traza sísmica, estos primeros arribos están asociados con la máxima concentración de energía de las ondas refractadas, y constituyen un reto para la geofísica en términos de precisión debido al ruido ligado a la calidad del proceso de adquisición que enmascara el primer arribo y dificulta la tarea de identificación. La detección de primeros arribos en trazas sísmicas es fundamental para numerosas aplicaciones en sísmica de prospección, ya que tiene utilidad en el procedimiento de tomografía sísmica y de correcciones estáticas.

Si bien, la detección de primeros arribos sísmicos se realizaba tradicionalmente mediante el picado manual por un reconocimiento visual de las amplitudes y variaciones de forma de onda

en los datos sísmicos, al contarse con grandes volúmenes de datos, la inspección visual de los registros se vuelve una tarea muy demandante en cuanto a tiempo y recurso humano, además que implica picados sesgados debido a la dependencia de la subjetividad del operador que realiza la labor. Actualmente, la gran mayoría de los instrumentos utilizados en las estaciones sísmicas digitalizan y almacenan los datos. El constante crecimiento del monitoreo sísmico ha motivado al desarrollo e implementación de nuevas y modernas soluciones automáticas con base en las detecciones manuales ya existentes, para efectuar la detección y picado de primeros arribos de manera automática o semi-automática.

En la exploración sísmica, el proceso de detección de primeros arribos, se ha efectuado mediante diversas técnicas cuyas características están enfocadas en aspectos probabilísticos, estadísticos y modelos matemáticos. Una correcta detección del primer arribo está condicionada por diversas propiedades al momento de la adquisición y exploración sísmica. Para la detección del primer arribo se han utilizado sistemas de procesamiento sísmico basados, por ejemplo, en relaciones de energía como el método de promedio de tiempo corto sobre promedio de tiempo largo (STA / LTA), descriptores de forma de onda más complejos Ruiz (1994); Allen (1978), la correlación cruzada de formas de onda (CCT) Mustafa Senkaya (2014) y la transformada de onda continua (CWT) Bogiatzis e Ishii (2015), donde la precisión al seleccionar el tiempo de llegada persiste debido al ruido de fondo, la variabilidad de las formas de onda sísmica y otros factores.

El método de Coppens corresponde a uno de los métodos ampliamente utilizados en la detección de primeros arribos, Coppens (1985) utiliza una función de energía que se aplica a las

trazas mediante una ventana deslizante. La función de energía calcula la suma de las amplitudes de las señales dentro de una pequeña ventana, que luego es normalizada por la suma de las amplitudes de todas las señales dentro de la ventana deslizante en la serie. El factor de normalización aumenta a medida que se calcula la función de energía para las señales posteriores, como resultado, el valor devuelto es más alto para el gran cambio de energía anterior que suele ser el primer arribo. No obstante, cada una de estas técnicas están asociadas a varios parámetros que, nuevamente, deben ajustarse manualmente y dependen en gran medida del tipo de datos disponibles para su estudio. Durante las últimas décadas, con el auge de las redes neuronales, se han desarrollado nuevas alternativas para la detección de primeros arribos (Ibarra, 2011), para facilitar y optimizar los procesos que no son tan claros para el ser humano.

En el presente trabajo se propone la detección de los primeros arribos en trazas sísmicas de sismogramas sintéticos generados con características del Valle Medio del Magdalena empleando un sistema basado en una red neuronal convolucional. Este proyecto se encuentra distribuido en seis capítulos. El Capítulo 1 se define el objetivo general y los objetivos específicos para el desarrollo del presente trabajo. El Capítulo 2, se definen los conceptos básicos dentro de la exploración sísmica con el propósito de comprender la teoría, se aclaran una serie de aspectos prácticos y se establecen las herramientas computacionales requeridas para el desarrollo de este proyecto. Debido a la gran diversidad de métodos para identificar los primeros arribos en trazas sísmicas, en el Capítulo 3 basados en la revisión del estado del arte, se exponen y comparan dos alternativas diferentes de solución, las cuales realizan la detección del primer arribo implementando dos topologías

diferentes de red: la primera se encuentra basada en dos redes neuronales densamente conectadas las cuales mediante el uso de diferentes transformadas (e.g. entropía), entregan información a cada red en ventanas antes y después del primer arribo; la segunda estrategia hace uso de una red neuronal convolucional que emplea 31 transformaciones Wavelet y el uso de filtro pasabajas. Los principales resultados en estas implementaciones fueron expuestos y presentados a lo largo del capítulo permitiendo seleccionar la arquitectura más adecuada, la cual es utilizada para el diseño de red solución. En el Capítulo 4, se plantea el diseño de la red solución a partir de los resultados obtenidos con las estrategias implementadas en el Capítulo 3 y múltiples experimentos en donde son seleccionados los parámetros e hiperparámetros que conforman a la red. En el Capítulo 5, se presentan las conclusiones, aspectos a considerar dentro de la detección del primer arribo. Finalmente, en el Capítulo 6, se encuentran las recomendaciones para futuros trabajos y/o los aportes por parte de los autores de este proyecto.

1. Objetivos

1.1. Objetivo General

Detectar automáticamente los primeros arribos de un conjunto de trazas sísmicas sintéticas terrestres del valle medio del Magdalena mediante el diseño de un sistema basado en redes neuronales artificiales.

1.2. Objetivos Específicos

- Definir una topología de red neuronal artificial para la detección automática de eventos sísmicos en señales unidimensionales.
- Entrenar una red neuronal artificial basada en la topología definida para la detección automática de primeros arribos.
- Evaluar el desempeño de la red entrenada para la detección automática de primeros arribos, empleando datos sintéticos terrestres del valle medio del Magdalena.

2. Marco de referencia

2.1. Contexto geofísico

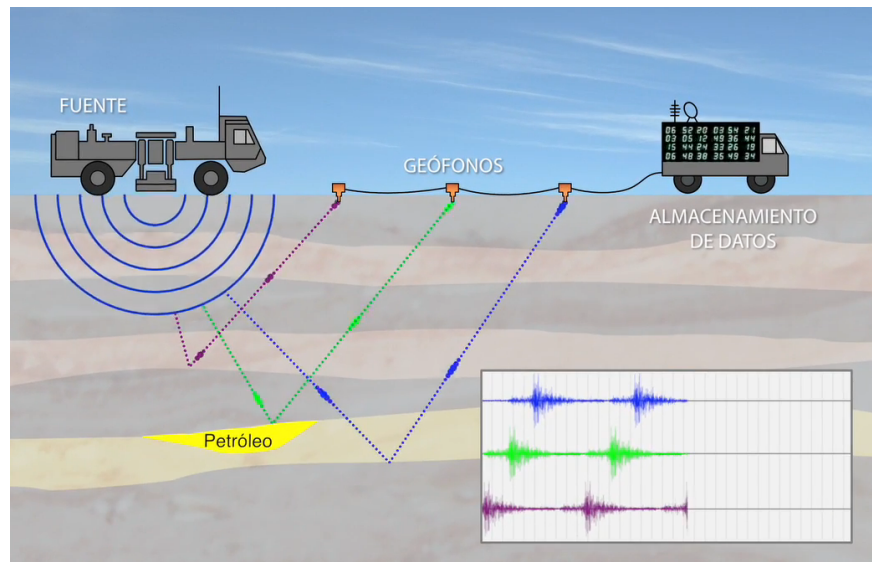
La geofísica es una de las principales ramas de las ciencias de la Tierra en la que se aplican conceptos de la física y la geología para el estudio del comportamiento de la tierra y se ocupa del estudio de las propiedades físicas de la Tierra. Incluye aspectos como la investigación de la composición interna del planeta, el flujo de calor del interior de la Tierra, la fuerza de gravedad que forma el campo gravitatorio, la fuerza de atracción magnética y la propagación de las ondas sísmicas a través de las capas de la corteza terrestre, (Krohn, 1984; Mexicano, sf). El estudio de la Tierra, es desarrollado a partir del análisis de los diversos fenómenos relacionados con la estructura terrestre, las condiciones físicas del planeta, así como de la historia de su formación y evolución. Esta ciencia usa métodos cuantitativos físicos, experimentos de observación, para recolectar información, interpretar, entender y predecir el comportamiento de nuestro planeta.

2.1.1. Método sísmico

El método sísmico se describe como el análisis de ondas sísmicas, las cuales son perturbaciones producidas en el subsuelo ya sean naturales o artificiales, las cuales se propagan a través de las diferentes capas del subsuelo y retornan a la superficie (Figura 1). El movimiento vibratorio en

forma de trenes de ondas concéntricas se propaga en diferentes direcciones a lo largo de cada capa. Las ondas experimentan reflexiones y refracciones cada vez que encuentran a su paso variaciones de velocidad y densidad entre capas. En estos cambios, parte de la energía es transmitida y otra reflejada. La energía que se refleja forma un campo de ondas que contiene información del medio por el cual viaja, brindando información sobre propiedades físicas, como la densidad, la elasticidad y la porosidad del medio (Lara, 2015). La velocidad de transmisión de estas ondas es constante para un mismo medio, pero varía en función de la rigidez (entre otros factores) del medio atravesado (Sabbione, 2012).

Figura 1. Método sísmico



Nota. Representación de adquisición de señales sísmicas en el contexto terrestre. Tomado de: Comisión Nacional de Hidrocarburos de México. (09 de septiembre de 2016). *CNH Video Autorizaciones Exploración Superficial* [Vídeo]. Youtube, <https://www.youtube.com/watch?v=8qeRvTpPHc>.

Las ondas sísmicas pueden generarse de forma natural, como ocurre cuando acontece un

sismo, o ser provocadas de forma artificial por el hombre al producirse una perturbación en el subsuelo mediante la aplicación de una fuerza externa. Para el caso artificial, se generan frentes de onda desde la superficie a partir del uso de unidades móviles que ocasionan vibraciones utilizando pequeñas cargas explosivas o fuentes mecánicas. Estas ondas viajan por el subsuelo atravesando capas de rocas. Los cambios de velocidad y densidad entre capas permiten que parte de la energía de la onda sea reflejada hacia la superficie. Cuando estas ondas llegan reflejadas a la superficie, son detectadas por unos sensores denominados geófonos para el caso terrestre e hidrófonos para el caso marino, la grabación del dato sísmico permite medir el tiempo transcurrido entre la emisión de las ondas y su llegada de nuevo a la superficie. Esta información se almacena de manera digital en cintas o discos duros en los equipos de registro del sismógrafo. La información sísmica recolectada se conoce con el nombre de traza sísmica o sismograma.

2.1.2. Onda elástica

Las ondas elásticas son perturbaciones realizadas a lo largo de un medio material en el cual se propaga de forma uniforme a través del mismo. La rapidez con que se propaga la onda elástica depende de las propiedades físicas (tales como el módulo elástico, la densidad, la temperatura, etc.) del medio material por el cual viaja la perturbación. El modelado de onda elástica representa un alto costo computacional. No obstante, el uso de herramientas como las GPU's facilitan la ejecución de modelos elásticos con mejores tiempos de ejecución y uso de memoria Kaus (2013). A partir de la ecuación de onda elástica, es posible simular el comportamiento de las ondas y adquirir información sobre el comportamiento de las ondas sísmicas durante su viaje sobre diferentes capas

terrestres. El modelado de ondas elásticas considera parámetros y coordenadas en el plano que describen la velocidad, presión, densidad y desplazamiento en las capas de la tierra.

2.1.3. Trazas sísmicas

Una traza sísmica representa la respuesta del frente de onda a la velocidad y los contrastes de densidad a lo largo de las interfases de las capas de rocas o los sedimentos a medida que la energía viaja a través de las capas del subsuelo, desde una fuente hasta un receptor o un arreglo de receptores (Barnes, 2007) (Figura 3). Las trazas sísmicas son usualmente almacenadas en archivos con formato SEG-Y.

Según el Servicio Geológico Colombiano (2002) el formato SEG-Y se rige por los estándares establecidos por la *Society of Exploration Geophysicists (SEG)* y contiene datos geofísicos en formato binario y textual, que incluyen la frecuencia de muestreo, tipo de disparo, línea de secuencia de rastreo, registro de campo, número de traza, código de identificación de la trazas, entre otros metadatos.

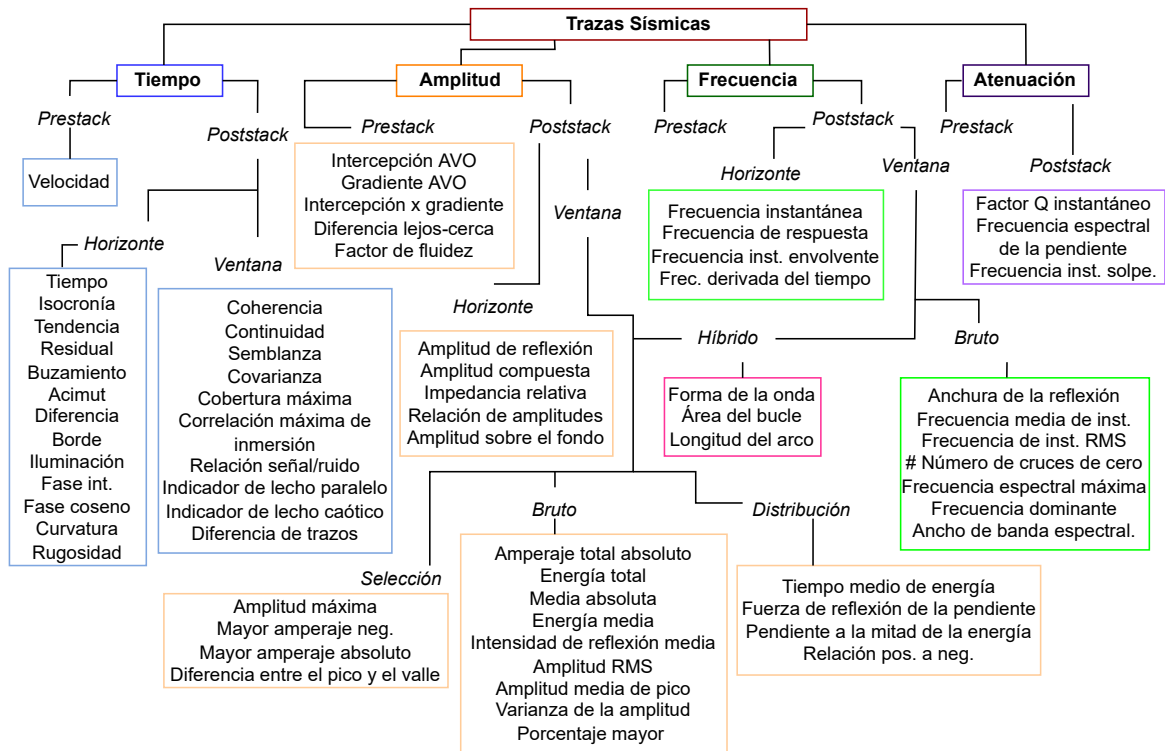
2.1.4. Atributos sísmicos

Un atributo sísmico es una medición de características geométricas, cinemáticas, dinámicas o estadísticas obtenida a partir de datos sísmicos, usualmente basada en mediciones de tiempo, amplitud, frecuencia, y/o atenuación (Barnes, 2016). Los atributos sísmicos permiten mejorar visualmente e inclusive cuantificar características de interés durante los procesos de interpretación

sísmica. Así, un atributo sísmico es adecuado en la medida en que sea sensible a una característica geológica determinada o que permita inferir propiedades de interés. Los atributos sísmicos son ampliamente utilizados actualmente en interpretaciones geológicas y geofísicas a partir de datos sísmicos, (Chopra, 2007). Las mediciones de atenuación usualmente tienen un alto grado de incertidumbre. Estas están usualmente basadas en datos apilados o migrados. Sin embargo, las basadas en datos antes de apilar son usadas generalmente en la determinación de velocidad de apilamiento, AVO (Variación de amplitud con la distancia/desplazamiento (AVA/AVO)) y otros atributos. Debido a la existencia de múltiples formas de agrupar la información sísmica, los atributos constituyen un conjunto abierto, y ya que se basan en algunos pocos tipos de medidas, los atributos generalmente no son independientes. Los atributos son útiles en la medida que se correlacionan con alguna cantidad física de interés, (Brown, 2011). La utilidad primaria de los atributos es que en ocasiones ayudan a ver características, relaciones y patrones que de otro modo no serían notados. Los atributos sísmicos se clasifican de acuerdo a su modelado matemático en *convencionales* y en *no convencionales*. Los diferentes atributos sísmicos permiten resaltar características de los datos, que contribuyan a la detección de los primeros arribos en señales sísmicas.

Los atributos geofísicos (Figura 2) registran propiedades de las ondas sísmicas y de las ondulaciones. Dentro del conjunto de atributos sísmicos, destacan: *la amplitud de la función envolvente*, *la fase instantánea*, *la frecuencia instantánea* y *la polaridad aparente*.

Figura 2.
Atributos sísmicos.



Nota. Atributos sísmicos, clases y derivaciones presentes en una traza. Figura adaptada bajo los principios descritos en (Brown, 2011).

2.1.5. Detección de primeros arribos

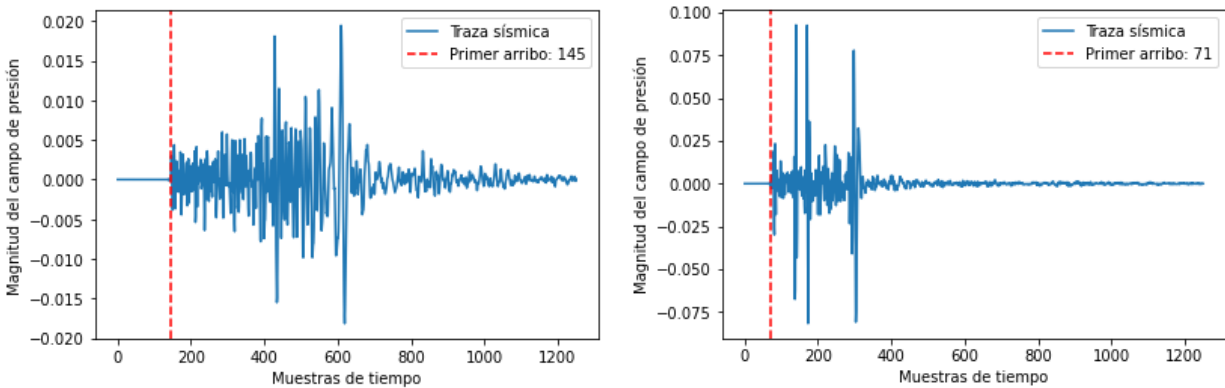
Dentro del concepto exploración sísmica, la detección o picado de primeros arribos (*first-breaks* o *first picking*)¹ consiste en determinar, dentro de un conjunto de trazas sísmicas, los tiempos de la onda primaria en cada traza con la mayor precisión posible. En general, estos arribos

¹ **FB:** First break - Primer arribo/ruptura.

están asociados con la energía de las ondas refractadas en la base de la capa no consolidada (capa de *weathering*) o con la energía de la onda que viaja directamente desde la fuente hasta el receptor (Sabbione, 2012).

Figura 3.

Representación gráfica de trazas sísmicas (trazas sintéticas con características del valle medio del Magdalena), generadas con su respectivo primer arribo.



(a) Traza sísmica con primer arribo en la muestra #145 (b) Traza sísmica con primer arribo en la muestra #71

Nota. Las trazas sísmicas de color azul pertenecen al conjunto de datos utilizados para el desarrollo del presente proyecto, dichos datos son crudos, lo cual indica que no cuentan con un procesamiento previo como lo puede ser una transformada, adición de ruido aleatorio u otra operación sobre la traza sísmica. El primer arribo en color rojo, pertenece al picado manual de la traza sísmica.

La detección del primer arribo parte de la identificación de cambios bruscos de amplitud en la señal sísmica, más precisamente, cuando en determinado instante de tiempo en la traza, es visible una variación en sus niveles de amplitud entre instantes de tiempo contiguos (Figura 3). La determinación precisa de los tiempos de primeros arribos es necesaria para el cálculo de

las correcciones estáticas², una etapa fundamental en el procesamiento de los datos sísmicos. La eficacia de las correcciones estáticas basadas en los métodos de reflexión y refracción depende de la fiabilidad del proceso de picado. Además, aplicaciones tales como correcciones estáticas por tomografía sísmica (*tomostatics*), requieren de una detección automática rápida y precisa de los primeros arribos de la señal.

2.2. Inteligencia artificial

La inteligencia artificial (IA) es la base a partir de la cual se imitan los procesos de inteligencia humana mediante la creación y la aplicación de algoritmos creados en un entorno dinámico de computación. El aprendizaje de máquina, también conocido como aprendizaje automático, es una rama de la inteligencia artificial que permite que las máquinas aprendan sin ser expresamente programadas para ello; una habilidad indispensable para hacer sistemas capaces de identificar patrones entre los datos con el objetivo de hacer predicciones sobre características observables en ellos. Todo proceso de *aprendizaje de máquina* (del inglés, machine learning) cuenta con dos etapas muy diferenciadas: entrenamiento y predicción. De acuerdo con los datos disponibles al momento de realizar el entrenamiento, el aprendizaje puede ser clasificado como *supervisado* o *no supervisado* (Larranaga *et al.*, 2007). Para el desarrollo de este proyecto se cuenta con un problema de aprendizaje supervisado de tipo regresión con el objetivo de identificar una relación funcional

² Las correcciones estáticas son corrimientos de tiempo a los datos sísmicos para llevarlos a un modelo idealizado en el que las fuentes y geófonos están ubicados en un nivel de referencia plano sin variaciones en elevación y velocidad, compensando las diferencias existentes en términos de topografía entre capas.

entre los parámetros de entrada y salida, debido a que los datos se encuentran etiquetados y el algoritmo predecirá una salida a partir de un rango continuo de valores posibles.

2.2.1. Algoritmos implementados para análisis de señales sísmicas

Habitualmente, la detección de los primeros arribos se desarrolla mediante el picado manual, el cual es llevado a cabo por un experto a partir de la inspección visual de los cambios de amplitudes y comportamientos de forma de la onda. Dentro de la detección manual, esta estrategia no sólo puede generar picados inconsistentes, ya que recae en la subjetividad del operador que realice la tarea, sino a su vez consume demasiado tiempo. Los diferentes niveles de complejidad geológica que puede presentar el subsuelo dentro del área de estudio, es la condición más rigurosa en términos de la calidad que se puede esperar para la detección de primeros arribos. Por tal motivo, el picado de los primeros arribos puede considerarse como una tarea ardua si los datos son adquiridos en un área con estructura subsuperficial no común, o si dentro de la adquisición se determina que la relación señal a ruido es baja. A su vez, es preciso aclarar que la calidad del picado del primer arribo depende del tipo de fuente, de la estructura del subsuelo más cercano a la superficie, con la relación señal-ruido de los datos y de la subjetividad del operador al momento de picar.

Durante las últimas décadas, se han realizado numerosas técnicas para determinar las primeras rupturas de forma automática o semiautomática. Los primeros intentos se basaban en el análisis expuesto por (Coppens, 1985). En este estudio, se introduce un criterio de comparación

de energías entre dos ventanas anidadas de diferentes tamaños. Otros enfoques se han centrado en la localización de las primeras llegadas a partir de la detección de cambios bruscos en la energía (Spagnolini, 1991). La metodología de las dos ventanas de tiempo mencionadas, son denominados métodos de tipo STA/LTA (*short term average / long term average*). El evento se declara cuando este cociente supera un valor de umbral (Allen, 1978).

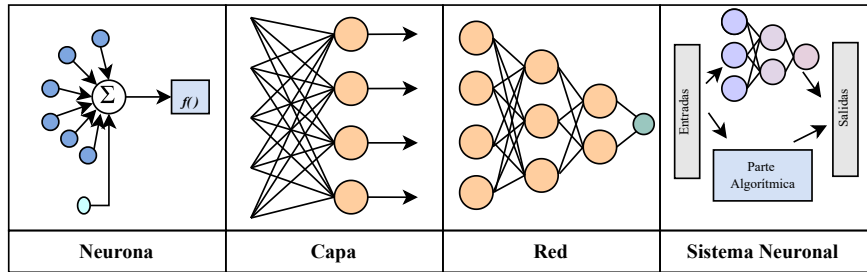
Algo semejante ocurre con el método presentado por (Earle y Shearer., 1994). En este método, a partir de la función analítica de la onda se define la envolvente de la señal como la función característica para obtener el cociente STA/LTA. Dentro de las propuestas para la selección del picado de primeros arribos, se han enunciado alternativas enfocadas en modelos matemáticos, como la teoría de fractales en sismica de exploración (Boschetti *et al.*, 1996), el uso de estadística de órdenes superiores (Saragiotis *et al.*, 2004; Yung, 1997), la aplicación estadística de segundo orden como la correlación cruzada (Clement y Peraldi, 1972; Shtivelman y Gelchinsky, 1983) y del uso de la transformación de señales en diferentes dominios, como la transformada Wavelet (Tibuleac *et al.*, 2003), como técnicas de visualización del comportamiento del subsuelo bajo diversas características.

2.2.1.1. Redes neuronales. Las redes neuronales artificiales son un paradigma de aprendizaje y procesamiento automático inspirado en la forma en que funciona el sistema nervioso de los humanos. Se trata de un sistema de interconexión de neuronas que colaboran entre sí para producir un estímulo de salida. Los elementos y la arquitectura de la red neuronal se describen en la Figura 4. Las redes neuronales artificiales (RNA) constituyen una herramienta útil que ha sido aplicada

a diversos campos del conocimiento. Existen diferentes tipos de las RNA (GAN³, CNN⁴, FNN⁵, RNN⁶, entre otras) de acuerdo con su aplicación en ingeniería y en geociencias.

Figura 4.

Estructura de una red neuronal



Nota. Se introduce el denominado modelo estándar de neurona artificial según los principios descritos en (Rumelhart *et al.*, 1986) y (McClelland y Rumelhart, 1986).

El propósito de las RNA es proporcionar una parametrización cuando se tiene un mapeo no lineal entre entradas y salida. La aplicabilidad del aprendizaje profundo y aprovechando el campo de reconocimiento de patrones, contribuyen al uso de redes neuronales para procesamiento e interpretación de datos sísmicos, como por ejemplo, la selección automática de la primera llegada mediante *Deep Learning* con aprendizaje interactivo humano (Tsai *et al.*, 2020), detector de eventos sísmicos en tiempo real utilizando redes neuronales (Barráez, 2000) y selección automática de

³ Las Redes Generativas Antagónicas (RGAs), también conocidas como (GANs) en inglés.

⁴ Red neuronal Convolutacional conocida por sus siglas en inglés (CNN).

⁵ Red neuronal Feedforward conocida por sus siglas en inglés (FNN).

⁶ Red neuronal Profunda conocidas por sus siglas en inglés (DNN).

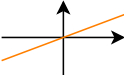

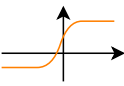

primeros arribos (Sabbione y Velis, 2010; Gelchinsky y Shtivelman, 1983), entre otros.

2.2.2. Función de Activación.

La función de activación entrega una salida a partir de un dato ingresado por el conjunto de entradas. Cada capa cuenta con una función de activación que le permite predecir la salida a partir del aprendizaje. Las diferentes clases de función de activación se describen a partir de la Figura 5.

Figura 5.

Funciones de activación.

Función de Activación	Ecuación	Gráfico
Lineal	$\phi(z) = z$	
Sigmoid	$\phi(z) = \frac{1}{1 + e^{-z}}$	
Tangente Hiperbólica	$\phi(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$	
Rectifier, ReLU (Rectified Linear Unit)	$\phi(z) = \max(0, z)$	

Nota. Representación gráfica y matemática de las diferentes funciones de activación en redes neuronales. Adaptado de: Stewart, M. (2019). *Intermediate Topics in Neural Networks, Towards Data Science*. Página web, <https://towardsdatascience.com/comprehensive-introduction-to-neural-network-architecture-c08c6d8e5d98>.

2.3. Entornos de trabajo

Para el desarrollo de este proyecto, las implementaciones, pruebas y revisiones de los prototipos, hasta el diseño, se emplearon las herramientas de hardware y software que se describen a

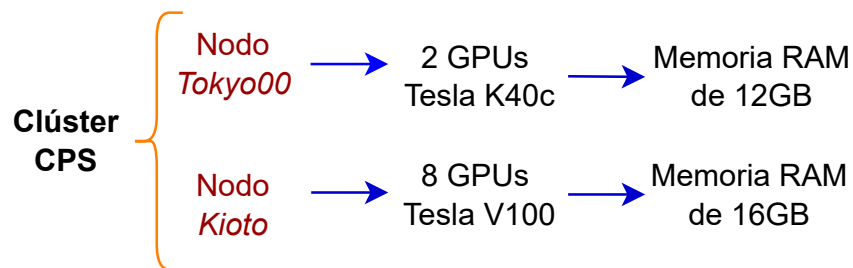
continuación.

2.3.1. Hardware

Los prototipos emulados del estado del arte y el modelo de red diseñado para el proceso de entrenamiento y validación de la red, se implementaron en Anaconda utilizando el clúster del Grupo de Investigación en Conectividad y Procesamiento de Señales (CPS). Los datos de los recursos computacionales con los cuales se contó para la ejecución de este proyecto se enumeran en la Figura 6.

Figura 6.

Características de recursos en Hardware



Nota. Durante las pruebas ejecutadas de los prototipos en el diseño de experimentos y de la red solución, se hizo uso de todos los núcleos disponibles en los nodos. Las pruebas e implementaciones se realizaron usando Keras.

2.3.2. Anaconda

La elaboración de los experimentos, prototipos y diseño de red solución, se realizó con la herramienta Anaconda para facilitar la construcción de los mismos. El código de programación está descrito en el lenguaje Python. Para los diversos procesos dentro del diseño de experimentos

y selección de la red, se utilizaron las librerías *segio*, *pywt*, *tensorflow*, *scipy*, *sklearn*, entre otras. A su vez, se utilizó Jupyter Notebook (Jupyter, 2022) y Google Colab (Google, sf), como entornos para pruebas preliminares.

Anaconda es un gestor de entornos de código abierto que abarca una serie de librerías y herramientas diseñadas para el desarrollo de la ciencia de datos con Python (Toro, 2022). Hemos seleccionado el uso del entorno en Anaconda, debido al control y organización de versiones de librerías y paquetes que ofrece, sin tener que configurar desde cero cada prueba que se desee ejecutar.

2.3.3. Librerías y paquetes

Dentro de la implementación de los diversos algoritmos, se emplearon diversas librerías tales como: *segio*, *ptwt*, *tensorflow* y *keras*.

- **Segio:** Es una biblioteca que permite el uso y tratamiento de datos sísmicos en formato SEG-Y. En python, permite seleccionar la información presente en los metadatos para su procesamiento en aplicaciones sísmicas, (Segio, 2022).
- **Ptwt:** Es el paquete de herramientas para la implementación de la transformada Wavelet de PyTorch. Las clases de TW son definidas por las propiedades que cada una de ellas poseen, entre las más importantes, se tiene: la función de escala, el número de momentos de desvanecimiento, el soporte compacto, la regularidad y la simetría. De acuerdo con las pro-

propiedades requeridas para su implementación, este paquete cuenta con la transformada rápida de Wavelet (fwt) implementada en *wavedec*, la transformada Wavelet en tiempo continuo (cwt), la transformada wavelet en el dominio discreto (dwt) y la descomposición de la transformada Wavelet vectorial y matricial *wavedec2(X,N,wname)*, entre otras herramientas de la transformada de ondícula (Wavelet, 2022).

- **Tensorflow:** Es una plataforma de código abierto para el aprendizaje automático, cuenta con múltiples herramientas, bibliotecas y recursos que permiten la aplicación del aprendizaje automático en procesos con grandes cantidades de datos, realiza operaciones a partir de *earrays* multidimensionales de datos que reciben el nombre de tensores. TensorFlow puede correr en múltiples CPU's y GPU's (Tensorflow, 2021).
- **Keras:** Es una biblioteca de código abierto de redes neuronales en Python. Permite la construcción de modelos y prototipos utilizando capas convolucionales y recurrentes (CodeTD, 2020).

3. Prototipos implementados para selección de la estrategia de detección

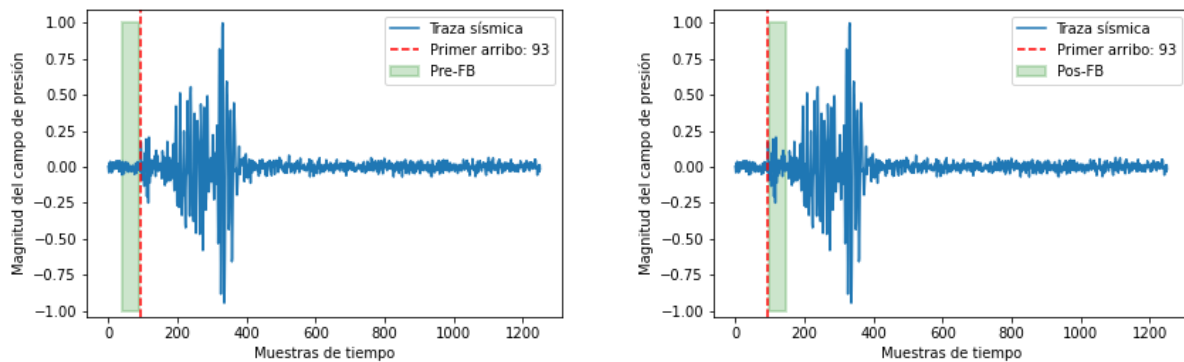
En este capítulo se presentan dos estrategias de solución para la detección de primeros arribos utilizando redes neuronales. La primera estrategia de detección del primer arribo por múltiples transformadas con el uso de una red neuronal profunda (DNN), está basada en el trabajo de Mężyk y Malinowski (2019). En este trabajo, los autores establecen y comparan tres alternativas para la detección de los primeros arribos empleando: entrenamiento de modelos generativos con redes neuronales densamente conectadas (DNN), regresión de vectores de soporte (SVR) y una implementación de árboles de decisión con aumento de gradiente (XGBoost, aumento de gradiente extremo), las cuales son comparadas entre sí para analizar los desempeños de las redes.

La estructura de solución planteada por el artículo está constituida por dos redes neuronales (DNN) con la misma arquitectura, como se muestra en la Figura 10. Una red neuronal es entrenada para realizar la detección del primer arribo presentándole una ventana antes del primer arribo (pre-FB) y la otra, se encarga de la detección del primer arribo siendo entrenada al presentarle una ventana posterior a él (pos-FB), en la Figura 7 se muestra el primer arribo de una traza sísmica con sus ventanas pre-FB y pos-FB. Para la realización de las pruebas, Mężyk y Malinowski (2019) definen la longitud de la ventana de 50 muestras y los parámetros utilizados en la red DNN, son presentados en la Tabla 2. Para encontrar el primer arribo, cada una de las DNN es entrenada a

partir de ventanas de muestras diferentes. La ventana de la DNN pre-FB se toma con una longitud de 50 muestras, la cual finaliza una muestra antes del primer arribo y para el caso de la DNN pos-FB, la ventana inicia en la muestra posterior al FB.

Figura 7.

Representación gráfica de la traza sísmica con su respectivo eventando antes y después del primer arribo.



(a) Traza sísmica con pre-Fb verdadero

(b) Traza sísmica con pos-Fb verdadero

Nota. El primer arribo ilustrado con una línea punteada de color rojo corresponde al valor del picado manual de la traza. A partir del picado manual el eventando pre-FB y pos-FB corresponden a las instancias verdaderas para la red pre-FB y pos-FB respectivamente. Los eventandos están constituidos por una longitud de 50 muestras cada uno (sin contener el primer arribo).

La segunda estrategia para la detección de primeros arribos hace uso de una CNN propuesta por Gillfeather *et al.* (2021), en la cual se realiza la comparación de tres tipos de redes neuronales para la predicción del tiempo de llegada del primer arribo en señales sísmicas, evaluados en un estudio de refracción sísmica de un depósito de mineral de hierro detrítico en Australia Occidental. Las redes neuronales evaluadas son: una red neuronal totalmente conectada (DNN),

una red neuronal convolucional (CNN) y un tipo de red neuronal recurrente denominada red de memoria a largo plazo (LSTM). De los resultados presentados dentro de la comparación de las tres redes neuronales, se decide implementar únicamente la red convolucional, debido a que, obtuvo un mejor desempeño y mayor precisión durante el entrenamiento y validación del modelo. Cada una de estas redes difiere significativamente en su arquitectura. Para el caso de estudio se ejecutó la arquitectura mostrada en la Figura 21. Gillfeather *et al.* (2021) sugiere que los datos que ingresan a la red correspondan a la traza cruda normalizada con ruido, la salida de la traza al pasar por un filtro pasa bajas y 31 transformadas Wavelet de la traza.

El conjunto de datos empleado en la implementación y evaluación de las dos propuestas y posterior diseño de la red solución, fue elaborado y proporcionado por el grupo de investigación Conectividad y Procesamiento de Señales (CPS). Las características del dataset utilizado son presentadas en la Tabla 1.

Los datos fueron adquiridos mediante el uso de ondas elásticas sintéticamente generadas.

Tabla 1

Conjunto de datos sísmicos en formato SEGY

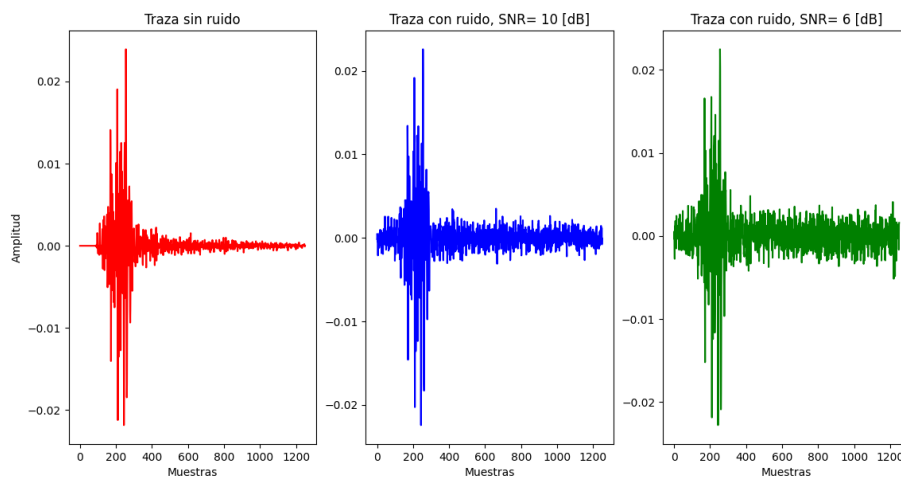
Especificaciones datos sísmicos en formato SEGY			
Nombre	Tamaño	N° de Trazas	Muestras por traza
SEAM_2d_DIP_ASIG_GEOM_planas_sou12_ rec06_off_3000_Local_sin_off0.segy	138,8 MB	27760	1251

Nota. El dataset corresponde a trazas sintéticas con características del Valle medio del Magdalena, las cuales fueron generadas sin ruido a partir de la ecuación de onda elástica. El conjunto de datos cuenta con una tasa de muestreo $T_s = 0,008$ segundos y una frecuencia de muestreo $F_s = 125$ muestras por segundo.

El conjunto de datos corresponden a trazas crudas generadas sintéticamente sin ruido de señales sísmicas, a las cuales se les realiza un preprocesamiento que consiste en la normalización de la señal sísmica y la adición de ruido gaussiano. Con lo anterior, el conjunto de datos de entrada corresponden a trazas crudas normalizadas con una relación señal a ruido (en inglés *Signal to noise Ratio*, abreviado SNR) variable de 6 a 10 dB,⁷ con el objetivo de presentarle a cada red un comportamiento más real y así optimizar su aprendizaje. La ejemplificación de este proceso, se visualiza en la Figura 8.

Figura 8.

Traza sísmica normalizada con adición de ruido



Nota. En la traza de color rojo se presenta la señal cruda, la traza de color azul contiene ruido gaussiano de 10dB y en la traza verde se aumenta el ruido gaussiano a 6dB. El ruido cuenta con una distribución uniforme a lo largo de toda la traza.

⁷ Este ruido se agrega de forma uniforme a lo largo de toda la traza. La SNR se incluye de forma aleatoria a cada una de las trazas dentro del rango nombrado.

Los datos crudos pasan por un preprocesamiento en el que la traza con ruido es normalizada y transformada. Los datos de entrenamiento y validación se normalizan con el propósito de comprimir las magnitudes del campo de presión para que estén en un rango definido, en nuestro caso, entre -1 y 1 . La normalización de las trazas es descrita por la Ecuación 1.

$$x_{nor} = \left(\frac{x_i}{\max(x_i)} \right) \quad (1)$$

Dónde x_{nor} corresponde a la traza normalizada, x_i es el i -ésimo valor en el conjunto de muestras de la traza, $\max(x)$ es el valor máximo entre todos los valores de la traza tomado de forma absoluta. Todos los datos utilizados para el desarrollo de este trabajo cuentan con ruido y normalización de -1 a 1 , esto también se cumple para las transformaciones realizadas en cada una de las implementaciones.

El desarrollo de estas implementaciones, se encuentra disponible para su revisión en el repositorio de GitLab ⁸ del Grupo de Investigación en Conectividad y Procesamiento de Señales (CPS).

⁸ Repositorio en GitLab del grupo CPS: https://gitlab.com/CPS_Repo/Seismic/semilleroan2020/fb-seismic.git

3.1. Primera estrategia: Detección de primeros arribos basada en múltiples transformadas

La selección automática del FB se realiza mediante la combinación de varios descriptores convencionales utilizados para caracterizar los inicios del FB, tales como el método de entropía. La selección del FB se plantea como un problema de clasificación binaria simple que requiere un modelo para diferenciar la muestra del FB de las que no son FB. Los datos de entrada son trazas crudas de señales sísmicas en conjunto con transformaciones de la traza original. Para el desarrollo de esta implementación se efectuó una de las 11 transformaciones propuestas en el artículo (Mężyk y Malinowski, 2019), esta consiste en la aplicación de la transformada de entropía con el objetivo que entregue características más relevantes y presentar una proyección en un dominio diferente para el entrenamiento de la red pre-FB y red pos-FB. La selección de esta transformada se ejecutó de forma arbitraria contando con la definición formal de la entropía en aplicaciones sísmicas, no obstante, cualquiera de las 11 transformaciones propuestas permitían su implementación.

La entropía se interpreta como la medida del desorden de un sistema en búsqueda del equilibrio. Cuanto mayor es la entropía de un sistema, menor es su disponibilidad de energía y mayor es el desorden del mismo (Córdova, sf). Denis y Crémoux (2002) calculan la entropía de una serie temporal en función del tiempo mediante la ecuación 2,

$$H(t) = \log \left(\frac{L(t)}{t} \right) \quad (2)$$

Donde $L(t)$ es la *longitud* de la serie temporal y se aproxima por la suma de los valores absolutos de

las primeras diferencias. Para calcular la entropía como atributo de la traza sísmica $s(t)$, se estima un $H(t)$ dentro de una ventana móvil de longitud fija n_h y se asigna su valor al último punto de la ventana. La función es definida por la Ecuación 3.

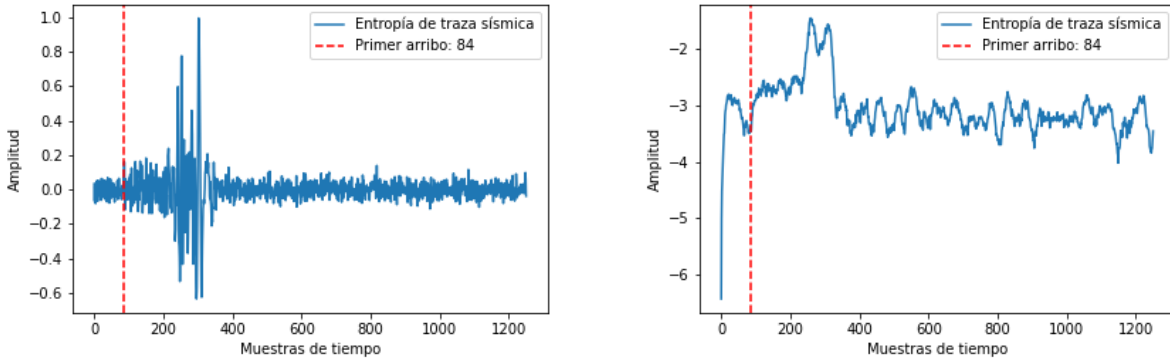
$$H(n) = \log \left(\frac{1}{n_h} \sum_{i=n-n_h+1}^{t-1} |s_{i+1} - s_i| \right) \quad (3)$$

En donde n_h corresponde a la longitud de la ventana, n el número de muestras de la traza, y s_i corresponde a las posiciones en muestras en cada una de las trazas. La representación gráfica de un cálculo de entropía es visualizada en la Figura 9.

Para el funcionamiento de la red neuronal pre-FB y pos-FB, se cuenta con un proceso en el cual, cada traza proporciona dos instancias, estas instancias corresponden a una muestra etiquetada con valor verdadero y otra con un etiquetado de valor falso (Figura 11). El etiquetado verdadero corresponde al picado realizado de forma manual en la traza y el falso es seleccionado de forma aleatoria en una muestra diferente a la verdadera.

Figura 9.

Entropía de una traza sísmica.

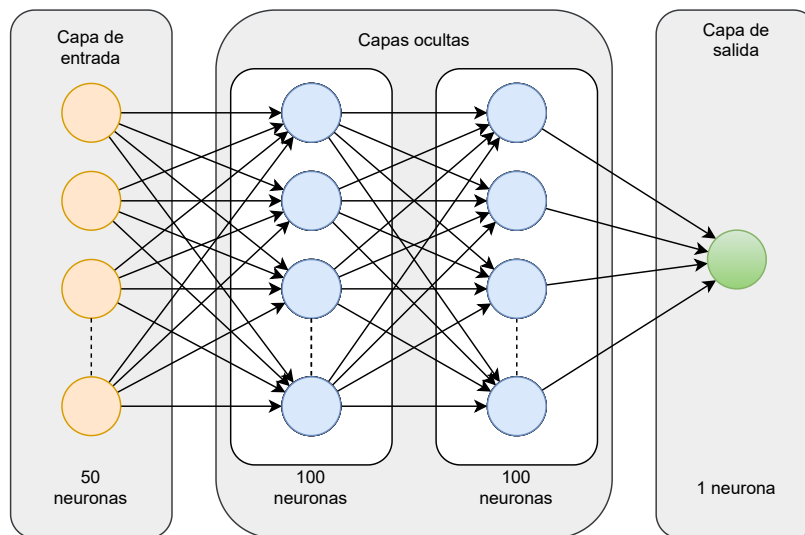


(a) Traza sísmica normalizada de -1 a 1.

(b) Transformada de entropía.

Figura 10.

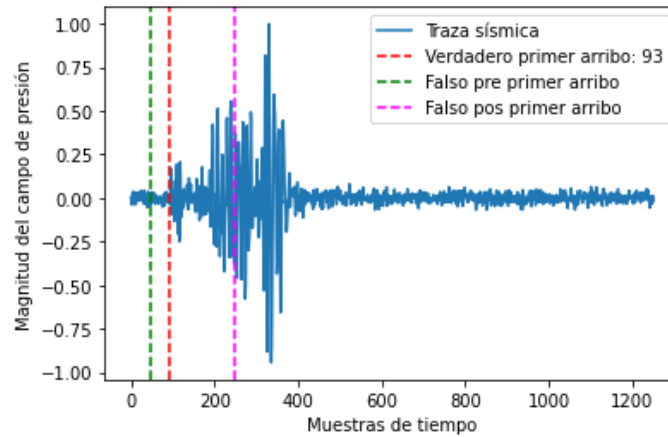
Arquitectura DNN para la red pre y pos primer arribo



Nota. Se implementan dos redes neuronales profundas para extraer características previas y posteriores al primer arribo. Las dos redes tienen la misma arquitectura, no obstante, al realizar el entrenamiento, difieren en la información que se presenta en cada red

Figura 11.

Primer arribo verdadero y falso.



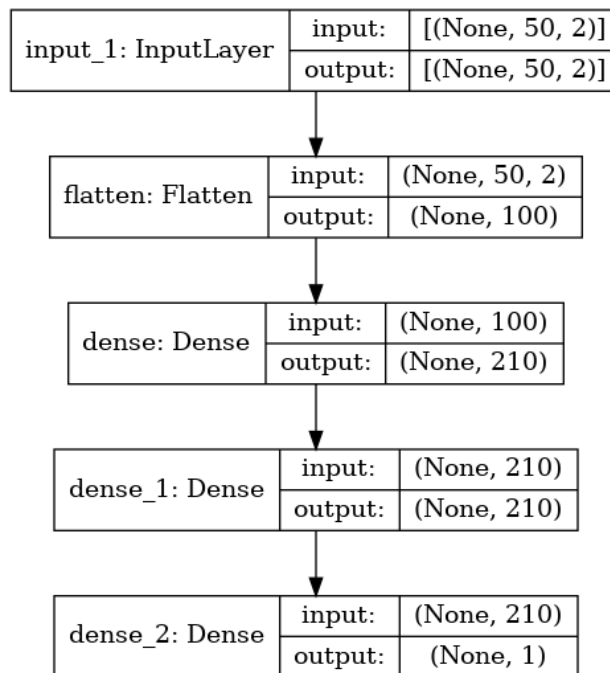
Nota. El primer arribo representado con color rojo, y los falsos primeros arribos representados con color verde y magenta, no se entregan a la red en una sola muestras como se presenta en la Figura 11, por lo contrario, se presenta una ventana de muestras por cada FB (tanto falso como verdadero) con 50 muestras, así como se ilustra por la Figura 14 y su correspondiente etiqueta.

A partir de la selección de muestras, tanto verdadera como falsa, son asignadas las ventanas para entrenar a cada una de las redes. En la red pre-FB, se le presenta una ventana que inicia antes de la instancia verdadera, esta ventana analiza las características de las muestras anteriores al primer arribo. A su vez, para la instancia falsa, se selecciona de igual forma una ventana, con la misma longitud, para las muestras anteriores a dicho FB. Estas dos ventanas (verdadera y falsa) se etiquetan con valor de 1 (para el caso verdadero) y 0 (para el caso falso) para ser presentadas a la red. De la misma manera, ocurre con la red pos-FB, una vez asignada las dos instancias (muestra verdadera y falsa) se asignan las etiquetas de 1 y 0 a cada una de las ventanas. El conjunto de datos de aprendizaje final, se construye mediante la recopilación de valores en eventos etiquetados como FB y no FB, repetidos para cada rastro sísmico.

Las instancias falsas tanto en la red pre-FB y pos-FB pueden tener valores en cualquier muestra de la traza, es decir, el falso FB, puede tomarse en muestras anteriores y/o posteriores del picado verdadero, entre un rango de la muestra 50 a la 1200, con la condición de que no puede tomar el primer arribo verdadero.

Figura 12.

Arquitectura DNN para la detección del primer arribo

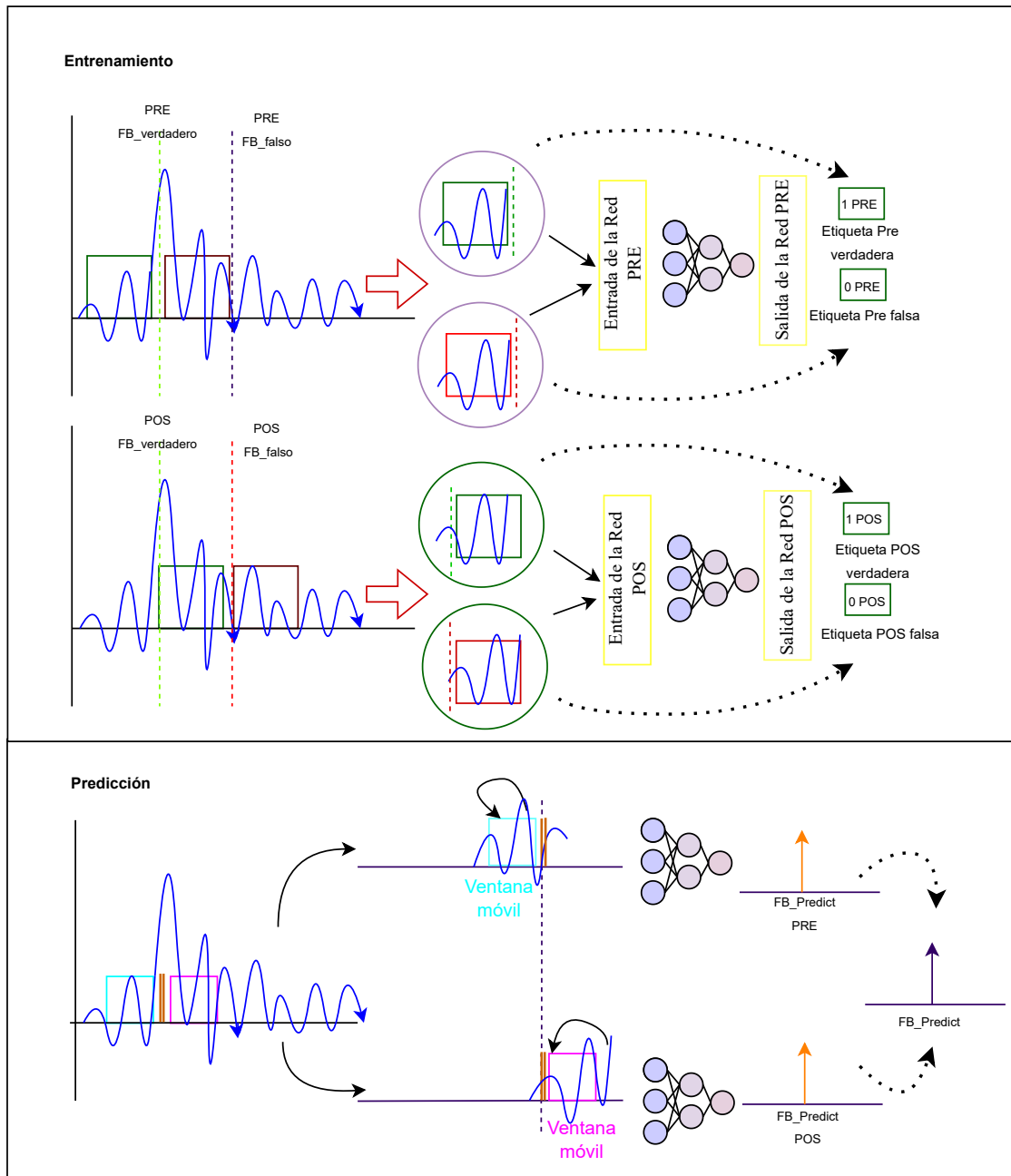


Nota. La red cuenta con dos capas ocultas, cada una de 210 neuronas. La entrada de la red esta constituida por 50 muestras de la traza cruda y 50 muestras de la transformada de entropía que corresponden a ventanas pre-FB o pos-FB, de acuerdo a la red que se desee entrenar.

Los datos de entrada a las redes están organizados en una matriz de características, la cual está constituida por ventanas de tiempo antes y/o después del evento FB (verdadero y falso), con el fin de poder entrenar la red pre-FB y pos-FB con etiquetados de cada traza sísmica ya

Figura 13.

Proceso realizado por las redes DNN pre y pos primer arribo

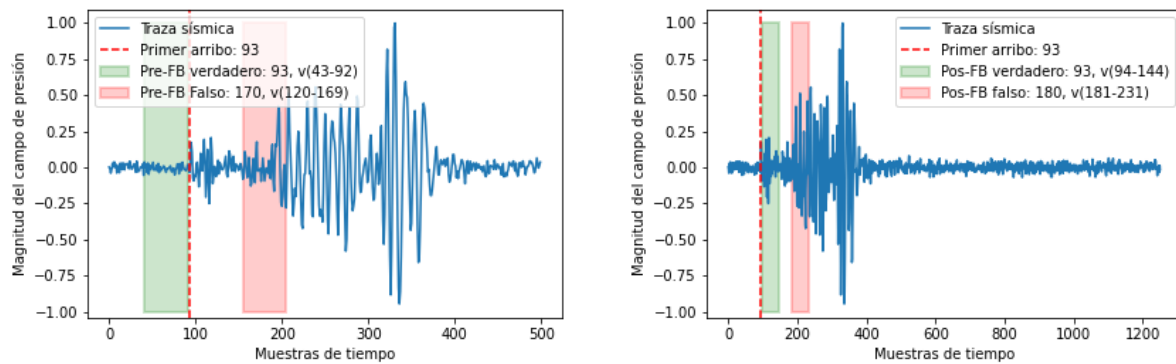


Nota. Representación del comportamiento de la red pre-FB y pos-FB. En entrenamiento cada traza con sus dos instancias, entrega a la red una etiqueta y un vector de muestras. Para predicción, la red buscará el patrón que correlaciona las dos etiquetas ya aprendidas, entregando un vector de probabilidades. A la salida, la muestra FB_Predict es el resultado obtenido de la intersección entre red pre y pos (multiplicar a FB_Predict PRE con FB_Predict POS). El enventanado móvil se realiza muestra a muestra.

mencionados. Como se mencionó anteriormente, se crean dos redes neuronales profundas (DNN) de iguales características, que son presentadas en la Tabla 2, el entrenamiento de las redes cuenta con un tiempo de ejecución de 300 épocas. Cada una de las DNN se encarga de encontrar el primer arribo, siendo éste la muestra que mayor peso tenga tras multiplicar los porcentajes de aciertos por muestra entre las dos DNN (red pre-FB y pos-FB).

Figura 14.

Enventanado pre y pos del primer arribo con etiqueta verdadera y falsa.



(a) Traza sísmica con etiqueta pre-FB verdadera y falsa. (b) Traza sísmica con etiqueta pos-FB verdadero y falso.

Nota. El primer arribo ilustrado con una línea punteada de color rojo, pertenece al FB de la traza mostrada en color azul. Esto ocurre de igual forma, para la red pos-FB. El término "v(43-92)" corresponde a la ventana definida en la muestra inicial de 43 y final en 92.

La arquitectura y el modelo definido para cada red es representado a partir de la Figura 10.

El conjunto de datos, contemplado en la Tabla 1 cuentan con trazas cuyos valores de primer arribo

oscilan entre la muestra 4 hasta la 148.⁹ Al contenerse valores de primeros arribos en muestras inferiores al tamaño de la ventana, decidimos utilizar únicamente las trazas cuyo primer arribo estuviese superior a la muestra 50. Esta selección de trazas dentro del conjunto de datos nos permitió contar con **20537 trazas** de las 27760 originales. De este subconjunto de trazas, **16429 trazas se emplearon para entrenamiento y 4108 para validación**. Es importante mencionar, que para la selección de la muestra etiquetada como falsa, dicho valor aleatorio sólo puede ser seleccionado dentro del rango de muestras comprendido entre la 50 hasta la 1200. Estas decisiones de trazas con picados superiores a 50 e instancias entre muestra 50 a 1200, son debido a que se quiso solventar de forma rápida el desbordamiento de muestras. De lo mencionado anteriormente, para las muestras cuyos valores de FB son menores de 50 muestras tanto verdadera como falsa, la ventana pre-FB tendrá posiciones nulas (la longitud no es de 50 muestras como se planteó), situación que se replica para el caso de instancias falsas en la red pos-FB o si es tomada en muestras superiores a 1201. Al ser longitudes de ventanas diferentes, se cuenta con problemas de dimensionamiento en las entradas de la red.

La restricción mencionada en el rango de FB falsos ocurre para las dos redes profundas. A su vez, recalamos, que la selección de este FB falso nunca podrá tomar la posición del picado que ha sido picado de forma manual como verdadero. Los datos de entrenamiento, fueron entregados a la red, en conjunto, es decir, ingresa una traza normalizada y su correspondiente transformada

⁹ En el archivo SEG-Y uno de los metadatos son los primeros arribos definidos en *ms*. Para este proyecto, se consideró el uso de muestras *m* como eje de tiempo.

Tabla 2

Definición de parámetros para la arquitectura DNN

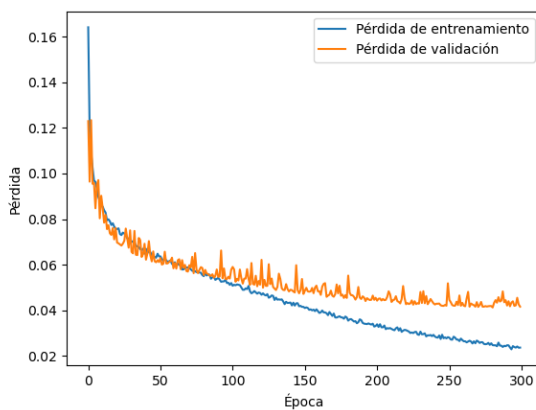
Función de Activación	Óptimizador	Inicializador	Batch size	Capas ocultas	Número de neuronas	Función de pérdida
Softsign	Nadam	Glorot uniform	470	2	420	MSE

Nota. Los parámetros fueron seleccionados a partir del resultado de la evaluación y comparación del desempeño de la red DNN.

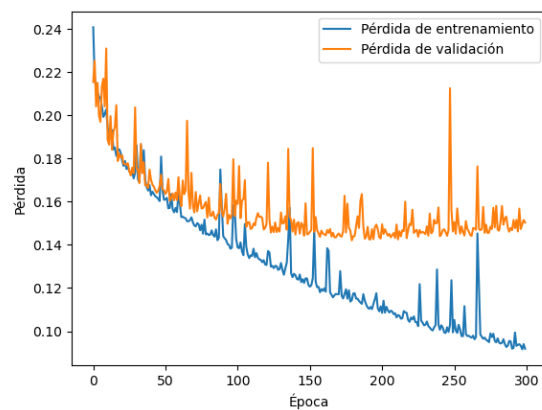
de entropía. La realización de las gráficas de función de pérdida y validación permiten identificar el nivel de aprendizaje que ha adquirido la red, el cual consiste en minimizar la diferencia entre el resultado predicho y el verdadero, determinando así el grado de convergencia, precisión y mejora progresiva del modelo al pasarle nuevos datos a lo largo de las épocas. Las curvas de la función de pérdida de entrenamiento y validación son mostradas en la Figura 15.

Figura 15.

Función de pérdidas de entrenamiento y validación en las redes pre-FB y pos-FB.



(a) Función de pérdida y validación para la red pre-FB.



(b) Función de pérdida y validación para la red pos-FB.

En la Figura 15.b, observamos que la red pos-FB presenta *overfitting* o sobreajuste. Esta respuesta ocurre por la poca variabilidad presente en el rango de SNR añadido. Al ser un intervalo de 6 a 10dB, el número de trazas que cuentan con niveles de ruido considerables aumentan, es decir, para cada nivel de ruido son distribuidas equitativamente el mismo número de trazas. Esto permite que a intervalos más pequeños, aumenta el número de trazas que cuentan con grandes proporciones de ruido por nivel, a comparación de tener un intervalo más grande de concentraciones de ruido pero con un número de trazas reducido por nivel, en la Figura 16 podemos observar el número de trazas para intervalos diferentes de SNR. En la Figura 20 se comprueba que ampliar el rango de ruido influye en las funciones de pérdida de la red.

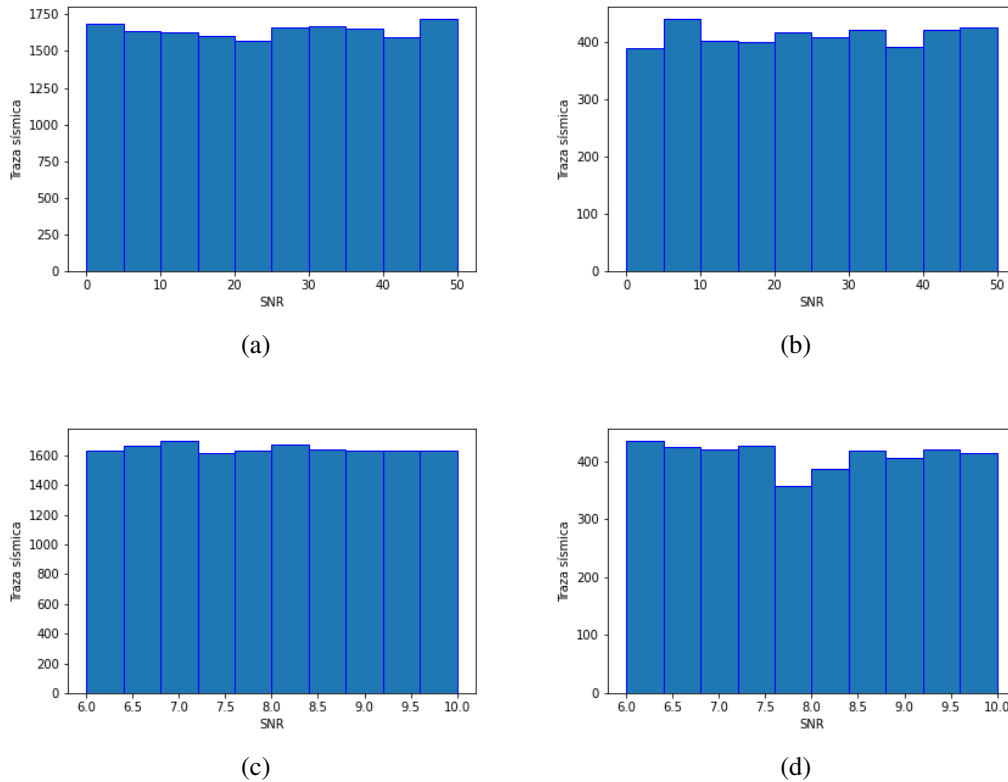
La selección del primer arribo realizada por las redes DNN ocurre bajo el criterio de intersección. Para el experimento de una sola traza, a la red ingresa la traza normalizada y su transformada de entropía con ruido. Cada una de las redes recibe las dos instancias con sus etiquetas. La salida, proporciona un vector de probabilidades en las que, de acuerdo a la función de costo, ajustará los pesos e indicará cuál de las 1251 muestras de la traza, cuenta con mayor probabilidad de ser el primer arribo. Dicho vector de probabilidades, es entregado por cada red tanto pre-FB como pos-FB. En cada una de las salidas de predicción se realiza un enventanado móvil el cual tiene una longitud de 50 muestras.

El enventanado está definido por medio de la Ecuación 4.

$$y(n+1) = \frac{1}{M} \sum_{m=0}^{m-1} x(n-m) \quad (4)$$

Figura 16.

Distribución de datos de entrenamiento y validación con diferentes niveles de SNR



Nota: El eje horizontal indica los intervalos de variación de la SNR y el eje vertical indica el número de trazas que cuenta con ese nivel de SNR. En (a) se presenta el histograma de distribución de trazas para entrenamiento con SNR de 0 a 50 dB, para (b) el histograma de distribución de trazas de validación con SNR de 0 a 50 dB, en (c) el histograma de distribución de trazas de entrenamiento con SNR de 6 a 10 dB y finalmente en (d) la distribución de datos de validación para una SNR de 6 a 10 dB. La red solución y sus diferentes implementaciones para su construcción contaron con trazas cuyo ruido añadido corresponde al intervalo de 6 a 10 dB.

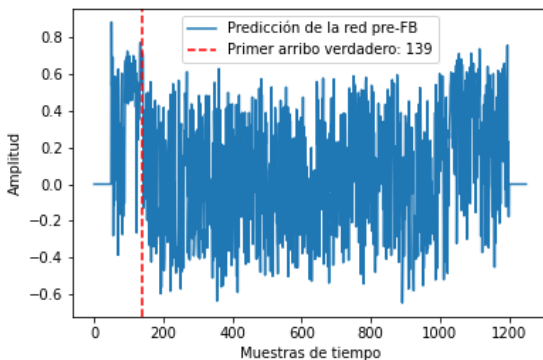
La primera posición del vector para la salida corresponderá al valor cuando $n=1251$, para cumplir con la dimensión del vector de salidas. La salida ideal de la red pre FB, consiste en una ventana con etiqueta de uno justo antes del primer arribo, para que al realizar el recorrido de la ventana móvil, el valor almacenado sea un 1 justo en la muestra del FB real. Lo mismo sucede para la red

pos FB, el resultado deberá ser 1 en la muestra donde el picado es verdadero. Una vez realizado este barrido para cada una de las redes, procedemos a multiplicar estos resultados. Las muestras en las que ambas redes se interceptan, entregan los intervalos en los que puede encontrarse el primer arribo, y tomando la máxima amplitud entre estos valores, se determina el primer arribo. Siguiendo el caso ideal, la multiplicación de las salidas, sólo será uno dentro de la muestra considerada FB verdadero.

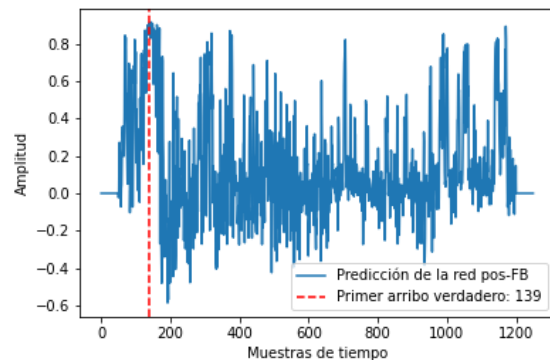
La Figura 13 muestra este procedimiento de predicción y la salida de la red pre-FB y pos-FB es mostrada por la Figura 17. Finalmente, la red entrega un valor de muestra predicho para la traza ilustrado en la Figura 18.

Figura 17.

Representación gráfica de salida para las redes pre y pos primer arribo.



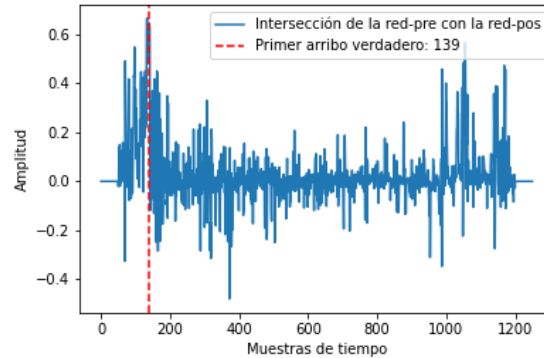
(a) Resultado del enventanado móvil en la salida de la red pre-primero arribo.



(b) Resultado de enventanado móvil en la salida de la red pos-primero arribo.

Figura 18.

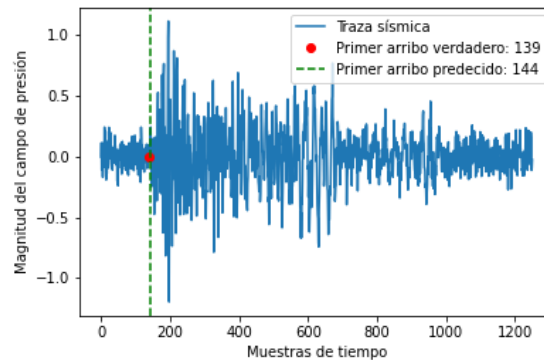
Identificación del primer arribo en la salida de las redes pre-FB y pos-FB.



Nota. En la figura se muestra el resultado de la multiplicación entre las salidas de la red. El primer arribo seleccionado por la red corresponde al máximo valor en amplitud de dicha multiplicación.

Figura 19.

Detección del primer arribo predicho por las redes pre y pos-FB.



Nota. La traza sísmica de color azul corresponde a la entrada de la red-pre y red-pos, que producen los resultados mostrados en la Figura 17, los cuales a su vez, por medio de la multiplicación, se obtiene la salida mostrada por la Figura 18, produciendo como resultado final el primer arribo representado con color verde en la Figura 19.

Durante la implementación de las dos redes DNN, el resultado de la multiplicación de los vectores de probabilidades proporcionados por cada red, permiten comprobar un buen desempeño

Tabla 3

Cálculo de métricas para las redes pre y pos FB con datos de entrenamiento y validación.

Datos	Media	Desviación estándar
Entrenamiento	66	173
Validación	65	170

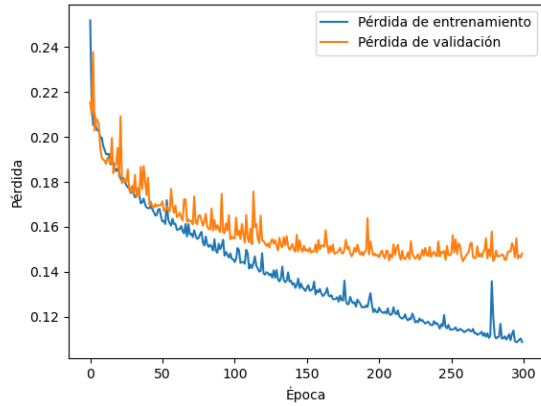
Nota. Para el caso de la media, se determina a partir de la resta entre el resultado del picado verdadero de los datos frente al picado predicho por la red.

en el proceso de predicción. En la Figura 19 podemos observar que la predicción del FB cuenta con una diferencia de cinco muestras corrida del valor verdadero, aun cuando la traza de entrada tiene grandes niveles de ruido, resultado que proporciona un desempeño. Sin embargo, también se presentan predicciones de arribos en muestras muy lejanas dentro del rango de FB del dataset, valores de FB en muestras como la muestra 1045 o predicciones máximas en la muestra 1099, cuando los arribos verdaderos oscilan entre la muestra 50 y 148. Estos grandes FB predichos son los que incrementan exponencialmente el error.

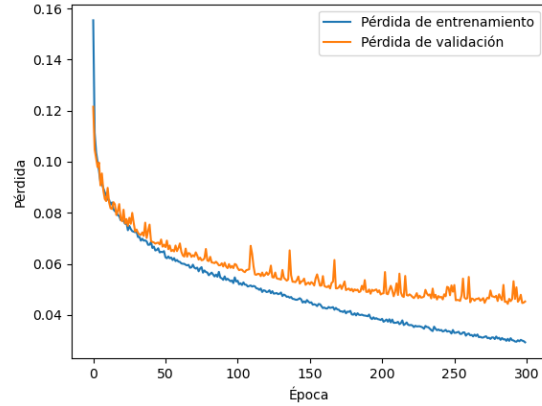
Dentro de la implementación de las dos redes DNN, decidimos aumentar el rango de ruido aleatorio agregado a las trazas, a un valor de 0 a 50 dB. De acuerdo a las Figuras 20, la función de pérdida mejora a lo largo de las épocas. Los resultados presentados en las gráficas de la Figura 20 alcanzan valores de pérdida de entrenamiento de 0.10 para el caso de la red pre y de 0.03 para la red pos primer arribo.

Figura 20.

Pérdidas de entrenamiento y validación por época, en la red pre-FB y pos-FB.



(a) Pérdida y validación para la red pre-FB



(b) Pérdida y validación para la red pos-FB

Sin embargo, este comportamiento no beneficia al entrenamiento de la red, ya que al ampliar el intervalo de ruido de 0 a 50dB aumentan los rangos posibles de ruido y se disminuye el número de trazas que posean un nivel de ruido considerable. Cada rango tiene una distribución uniforme lo que implica que el número de trazas en cada uno de ellos sea el mismo para todos. Las distribuciones de las trazas para cada nivel de ruido son representadas por los histogramas de la Figura 16. Al tener un ruido de 6 a 10dB, se concentra un mayor número de trazas en rangos con grandes cantidades de ruido. Al realizar el entrenamiento de la red con trazas menos ruidosas, el aprendizaje es más sencillo y rápido, ya que cada red identifica las variaciones de amplitud que ocurren en niveles diferentes de cero y aprende que después de un nivel de cero en amplitud, la siguiente variación notoria consiste en el primer arribo. Situación que no ocurre en la adquisición de señales sísmicas, ya que el FB muchas veces es enmascarado por el ruido presente al momento

de realizar el experimento sísmico.

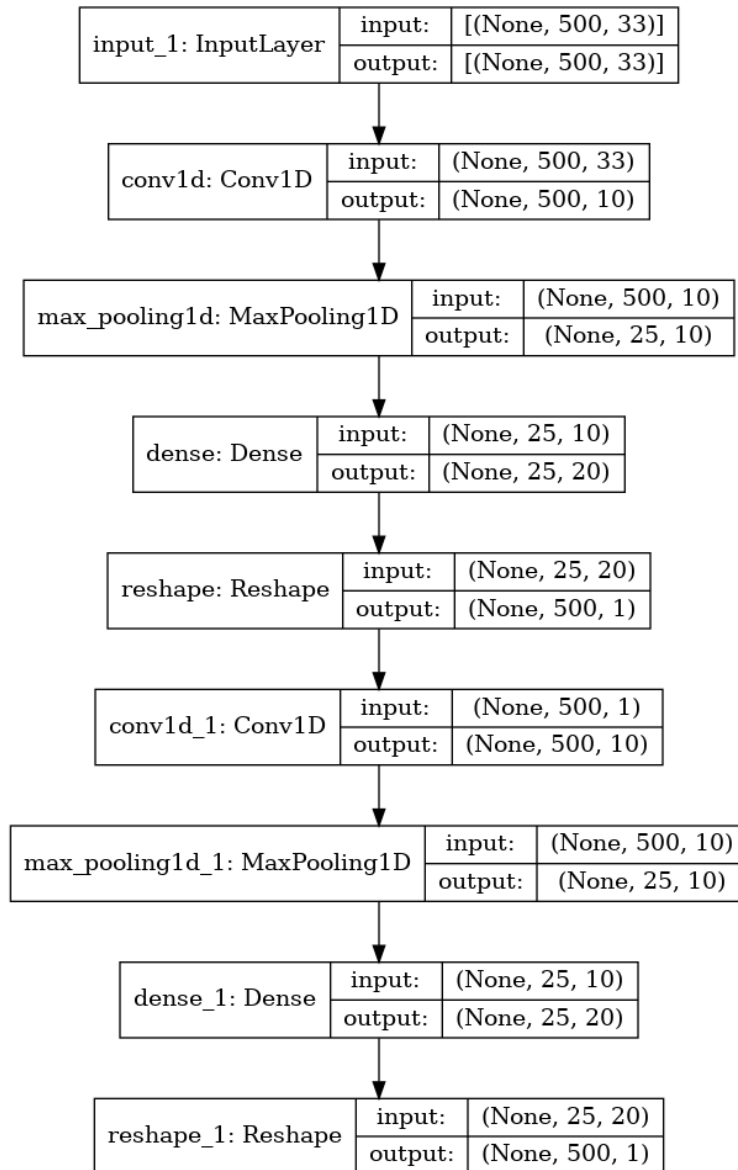
3.2. Segunda estrategia: Detección de primeros arribos basada en transformadas Wavelets

Para la segunda estrategia, la detección de primeros arribos a partir de redes neuronales, se realiza mediante el uso de una red neuronal convolucional (CNN). La señal de entrada se convoluciona con un *kernel* o filtro mediante una ventana de convolución deslizante. Durante el proceso de entrenamiento, cada nodo extrae de los filtros resultados unidimensionales de patrones temporales. Para una capa convolucional dada, cada nodo (filtro) es un kernel independiente de pesos con una longitud específica, estos nodos toman todas las entradas de una capa determinada y las convoluciona con el núcleo, los pesos del kernel para cada nodo distinto se inicializan aleatoriamente y se optimizan a través del proceso de retropropagación. De acuerdo con el artículo Gillfeather *et al.* (2021), se utilizan 10 filtros (kernel) en cada capa convolucional con un tamaño de 20 muestras, tomando un número n de entradas.

Cada filtro genera un conjunto de 500 valores, por lo que con los 10 filtros de la capa obtenemos una matriz de dimensión (500, 10). En la capa de *maxpooling*, se ajustan los datos de la entrada convolucional tomando los máximos valores de características dentro del tamaño del kernel. Las capas densamente conectadas se ubican con el propósito de generar nuevos pesos sobre el entrenamiento y adicionalmente, realizar el proceso de clasificación de las muestras, para entregar una salida con 500 muestras en conjunto con la muestra del FB ideal, que ha sido etiquetada a partir del uso de la función escalón unitario.

Figura 21.

Arquitectura CNN para la red de detección del primer arribo



Nota. La red cuenta con una capa convolucional, seguida de una max pooling, en conjunto de una Densa con Reshape, la mitad de la red es replicada para entregar una sola capa de salida. Este procedimiento se realiza con el objetivo de afianzar el modelo de aprendizaje siendo una red robusta para la detección del FB. La arquitectura es tomada con base en el artículo.

Para esta estrategia, los datos presentados a la red consisten en 20537 trazas del conjunto de datos mostrado en la Tabla 4, en donde se cuenta de igual forma que la red DNN con 16429 trazas para entrenamiento y 4108 trazas para validación, con la diferencia de tener una longitud de 500 muestras de las 1251 originales que contiene cada traza. Esta adecuación se realiza para implementar la red como es descrita en el artículo. Cada traza cuenta con el valor del picado manual (verdadero), lo que corresponde al mismo conjunto de datos implementado para la red DNN, con el objetivo de comparar los resultados para cada topología bajo las mismas condiciones.

Tabla 4

Conjunto de datos sísmicos para red CNN

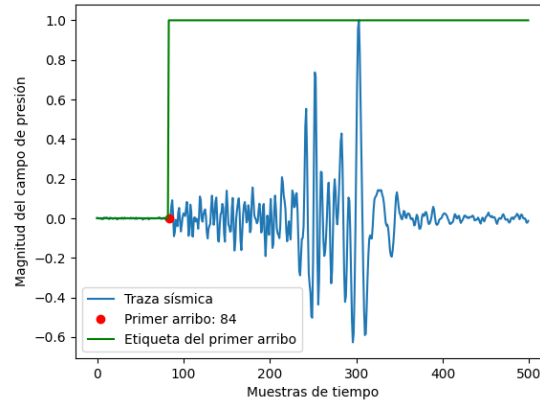
N° de trazas	Entrenamiento	Validación	Muestras por traza
20537	16429	4108	500

Nota. El conjunto de datos corresponde al set mostrado en la Tabla 1, las cuales fueron adoptadas a una longitud de 500 muestras para la implementación de la red CNN. El conjunto de datos cuenta con una tasa de muestreo $T_s = 0,008$ segundos y una frecuencia de muestreo $F_s = 125$ muestras por segundo.

Durante el entrenamiento de la red, el etiquetado se realiza mediante la función escalón unitario, el cual enseña a la red neuronal cuál debería ser la salida correcta, o ideal, para la entrada de una traza. Para el funcionamiento de la red neuronal convolucional, durante el entrenamiento y el aprendizaje de la red, la muestra en donde se encuentra el primer arribo verdadero (dato conocido por el picado manual) toma un valor de uno en todas las muestras posteriores al primer arribo, incluida la muestra en la que está ubicado. Para las muestras predecesoras al FB cuentan con una máscara cuyo valor es de cero.

Figura 22.

Etiqueta de entrenamiento en la red convolucional



Nota. El etiquetado (máscara del primer arribo) del entrenamiento consiste en asignar a las muestras antes del FB un valor de 0 y las muestras siguientes con un valor de 1 incluyendo el primer arribo.

La salida de la red consta de una sola capa, la cual entrega la muestra correspondiente a la detección del primer arribo. La CNN toma la salida ideal y adapta los pesos para obtener una salida más precisa la próxima vez que reciba una traza similar. Este proceso se repite con el objetivo, que en cada una de las épocas, se reajusten los pesos, para que la red tienda a seguir el comportamiento del escalón y así, seleccionar la muestra adecuada. Los parámetros utilizados para la implementación de la red CNN, fueron tomados de (Gillfeather *et al.*, 2021) y mostrados en la Tabla 5. Para la aplicación del filtro pasa bajas, se parte de una frecuencia de corte $f_c = 0.3$ determinada a partir del uso de la transformada rápida de Fourier, Figura 23. El cálculo de la frecuencia de corte es seleccionada mediante el uso de las frecuencias de operación en las que se concentran los primeros arribos de nuestro conjunto de datos.

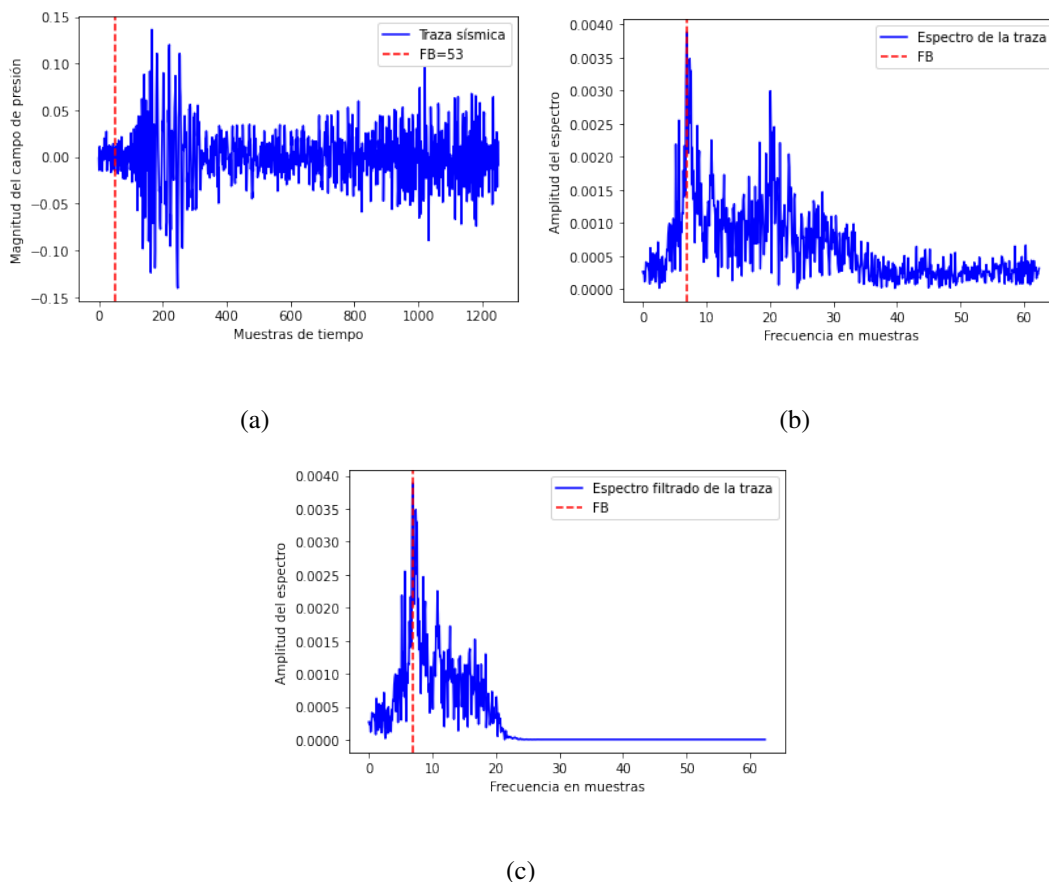
Tabla 5

Definición de parámetros para la arquitectura CNN

Función de Activación	Óptimizador	Métrica	Batch size	Capas	Función de pérdida
Tangente hiperbólica	Adamax	Binary_accuracy	20	8	MSE

Nota. Los parámetros fueron seleccionados a partir del resultado de la evaluación de desempeño para cada una de las 3 perspectivas de implementación presentes en el artículo Gillfeather et al. (2021). El número de capas corresponden a 2 capas convolucionales, 2 capas densamente conectadas, 2 maxpooling y 2 reshape.

Figura 23. Respuesta en frecuencia filtro pasa baja



Nota. Durante el entrenamiento de la red, a las trazas se les realiza un filtrado para eliminar las altas frecuencias. La frecuencia de corte fue determinada a partir de la *fft* comprobando que el filtro cumple para el rango de FB presentes en el conjunto de datos. En (a) se representa la traza sísmica con su primer arribo, en (b) la transformada de Fourier para la mitad de la frecuencia y en (c) se muestra la traza filtrada.

En correspondencia con lo presentado en el artículo, la red CNN cuenta con 31 transformadas Wavelets con el objetivo de entregar datos a la red a partir de diferentes representaciones de la traza y de esta forma adquirir más condiciones para encontrar el primer arribo. La transformada Wavelet o también conocida como la transformada de ondícula, descrita en Sánchez (2006), es un tipo especial de transformada matemática que representa una señal en términos de versiones trasladadas y dilatadas de una onda finita (denominada ondícula madre). La transformada de wavelet, fue determinada a partir de la librería *PyWavelets* y el método *pywt.cwt* en Python. La Wavelet continua (CWT) es la suma sobre todos los tiempos de versiones desplazados y escalados de la ψ (función Wavelet). La función Wavelet hace referencia al conjugado de la wavelet madre que será escalada y trasladada punto a punto para determinar los niveles de comparación con la señal *CWTF*. El CWT, cuando se aplica sobre la señal original $f(t)$, se expresa como,

$$CWTf(\tau, s) = \int_{-\infty}^{\infty} f(t) \psi_{\tau, s}^*(t) dt \quad (5)$$

El cálculo de la CWT a través de una señal nos da resultados en muchos coeficientes, que son funciones del parámetro de traslación τ y el parámetro de escala s .

$$\psi_{\tau, s}(t) = \frac{1}{\sqrt{|s|}} \psi\left(\frac{t - \tau}{s}\right) \quad (6)$$

De hecho, τ es proporcional a la información de tiempo. Se especifica la ubicación de la Wavelet en el tiempo; τ por la variación de la Wavelet puede ser desplazado sobre la señal. La escala s

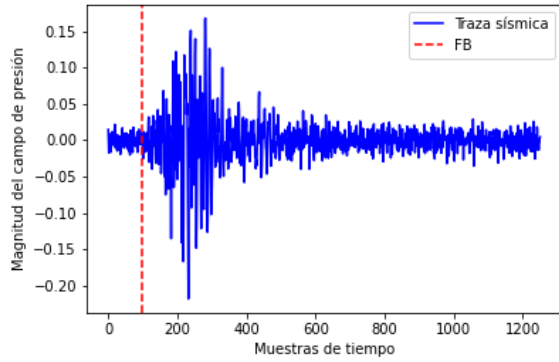
es inversamente proporcional a la frecuencia de centra de la información, s entrega la escala o dilatación de la Wavelet. La variación de s modifica no sólo la frecuencia central de la Wavelet, sino también la longitud de la ventana. Grandes escalas están relacionadas con las frecuencias bajas, dando la información global de la señal. Considerando que, escalas pequeñas corresponden a altas frecuencias, revelando los detalles más finos de la señal. Por otra parte, la energía de la señal se mantiene constante en cada escala, ya que se normaliza por $\frac{1}{\sqrt{|s|}}$.

Por lo general, el cálculo de la CWT se lleva a cabo mediante la adopción de valores discretos para el parámetro de escala s y el parámetro τ de traslación. Para su uso, se debe definir una variable, en nuestro caso *coefs* la cual nos indica el número de transformaciones requeridas. A su vez, es necesario determinar la ondícula madre a utilizar. Para la práctica aquí propuesta se utilizó la ondícula Morlet (*morl*). La ondícula Morlet, se encuentra definida por la Ecuación 7 y es aplicada por su distribución gaussiana ya que permite concentrar la mayor probabilidad dentro del arribo verdadero.

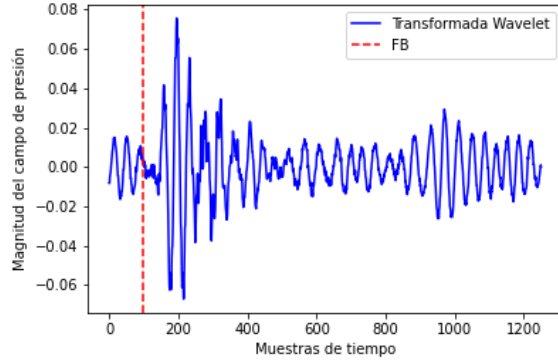
$$\psi(x) = e^{-\frac{x^2}{2}} \cos(5x) \quad (7)$$

Figura 24.

Traza sísmica con una de sus componentes Wavelet



(a) Traza n°2450 del dataset con un ruido de 10dB y con FB en la muestra 98.

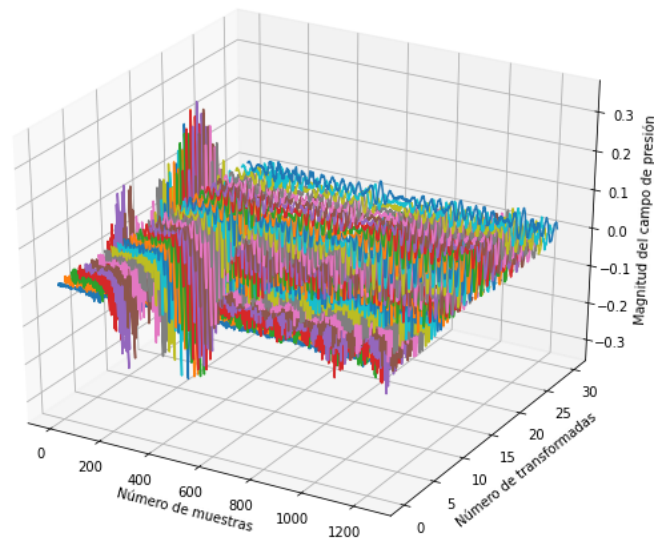


(b) Transformada wavelet n°31.

Nota. La línea roja punteada corresponde a la ubicación del primer arribo, picado de forma manual, para la traza.

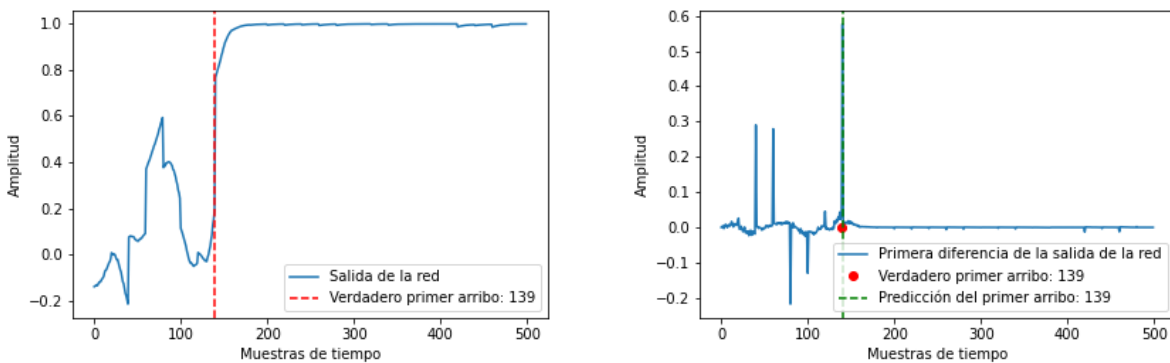
Figura 25.

Conjunto de transformadas Wavelet para una traza sísmica, red CNN



La selección del primer arribo realizada por la red ocurre bajo el criterio de primera diferencia. La primera diferencia, se realiza con el fin de obtener la máxima diferencia que existe dentro de la salida de predicción. Durante el proceso de entrenamiento, el etiquetado del FB se realiza utilizando un escalón unitario. Idealmente, se desea que la red entregue el mismo escalón unitario a su salida, y que al ingresar un dato nuevo, el ajuste de los pesos a lo largo de las capas, entregue un escalón con la ubicación de dicho FB. A partir del uso de las primeras diferencias, se está detectando la máxima variación que ocurre en la señal. Si la salida fuese un escalón ideal, la mayor variación ocurre justo en el instante en que la amplitud del escalón cambia de 0 a 1. Para la implementación de la CNN, la salida de predicción de la red es representada por la Figura 26 (a).

Figura 26. Salida de la red CNN



(a) Salida de la red para predicción CNN

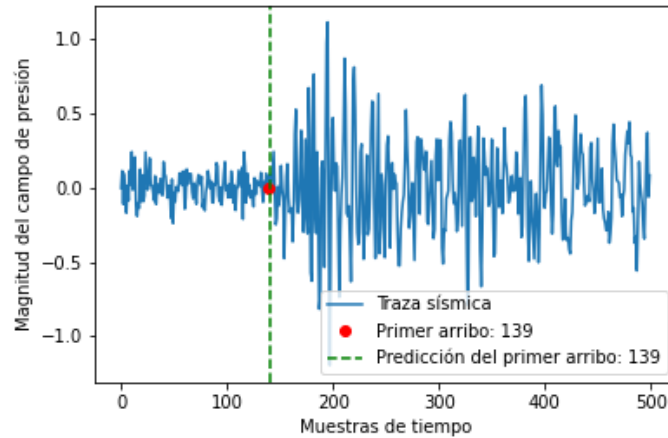
(b) Resultado primera diferencia con detección del FB.

Nota. Para la Figura 26(b) El primer arribo está representado por la máxima amplitud de la primera diferencia. Los resultados presentados son a partir de datos de validación, los cuales corresponden a datos que no habían sido conocidos por la red en la etapa de entrenamiento.

Una vez realizada la primera diferencia, se selecciona el pico máximo, el cual corresponde

Figura 27.

Estimación del valor del primer arribo con la red CNN

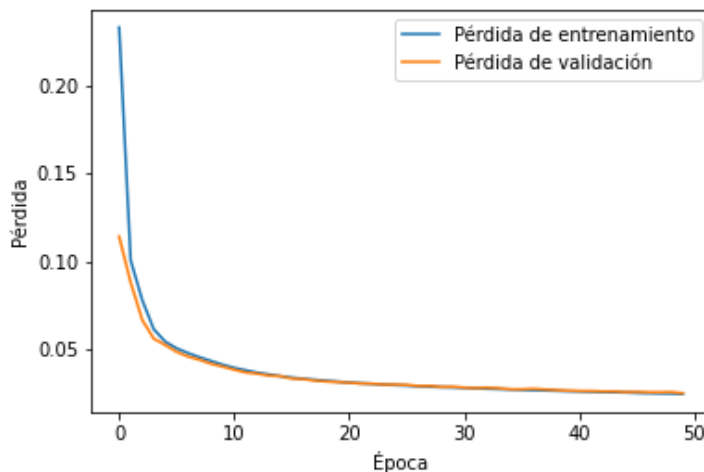


Nota. La predicción del primer arribo predicho se sobrepone en la muestra correspondiente a al primer arribo de la traza, evidenciando el bajo error de detección con la red CNN.

a la mayor variación entre muestras, presentada durante la salida de la red, por tanto, corresponde al FB. Finalmente, el valor que se predijo por la implementación de la red, es visualizado por medio de la Figura 27. A partir de la realización del entrenamiento de la red CNN, se obtienen las gráficas correspondientes a la función de pérdida para cada conjunto de datos: entrenamiento y validación, presentadas en la Figura 28.

Figura 28.

Función de pérdida de entrenamiento y validación de la red CNN



Nota. En la figura, se visualiza el aprendizaje de la red para 50 épocas. La curva de pérdida en el conjunto de validación inicia en un valor notoriamente bajo comparado con la de entrenamiento.

Con los resultados obtenidos en la predicción para el conjunto de 4108 trazas, se calcularon diferentes métricas para evaluar el desempeño de la red. Los cálculos realizados están contemplados en la Tabla 6.

Tabla 6

Cálculo de errores con datos de entrenamiento y validación.

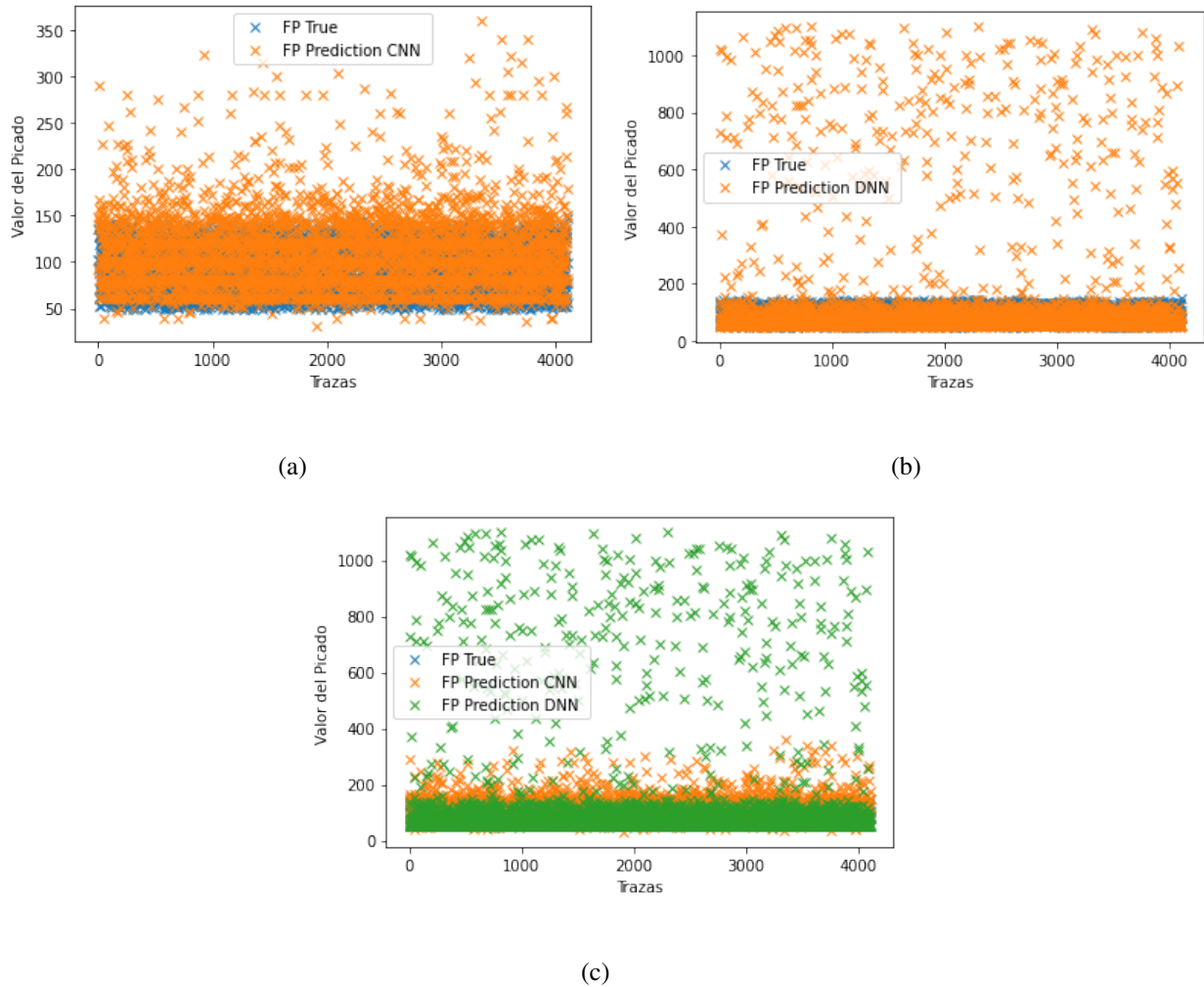
Datos	Media	Desviación estándar
Entrenamiento	54	93
Validación	54	91

Nota. Las métricas fueron calculadas en términos de muestras. Para el caso de la media, se determinó como el error absoluto entre el resultado del picado verdadero de los datos frente al picado predicho por la red.

3.3. Selección de Topología

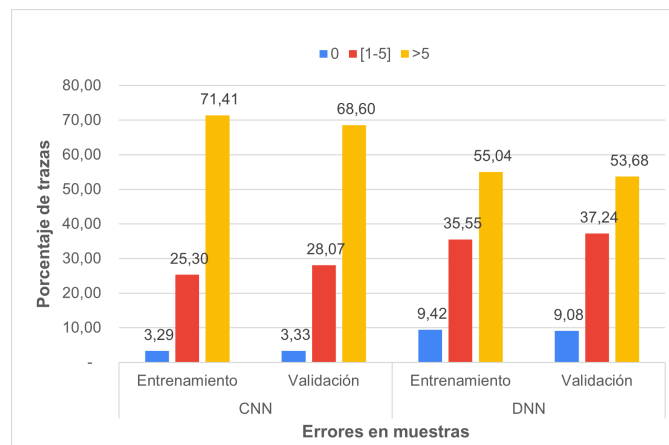
De acuerdo a la implementación de las dos estrategias, la red DNN y la red CNN, propuestas por los artículos (Mężyk y Malinowski, 2019; Gillfeather *et al.*, 2021), se concluye que si bien, ambas topologías realizan la detección del primer arribo de una traza sísmica y entregan resultados aceptables y cercanos al picado verdadero, la red CNN cuenta con mejor desempeño frente a la DNN. La CNN tiene una media de error ligeramente menor que la DNN pero los arribos incorrectos que detecta no sobrepasa el umbral de las 400 muestras, es decir, el rango de FB predichos por la CNN están entre la muestra 31 a la 360, caso contrario a la DNN en el que se detectan arribos en muestras entre la 49 y superiores a la 1099 (Figura 29), situación que no ocurre para los datos utilizados en el entrenamiento de las redes.

Figura 29. Distribución de predicciones de primeros arribos para las redes CNN y DNN



Nota. El conjunto de datos corresponden a 4108 trazas de las cuales en (a) se representan los primeros arribos predichos por la red CNN vs primeros arribos verdaderos, en (b) primeros arribos predichos por la red DNN vs primeros arribos verdaderos y en (c) Primeros arribos verdaderos comparados con las predicciones de las redes DNN y CNN.

Figura 30. Histograma con porcentajes de error en la predicción de primeros arribos para las redes CNN y DNN



Nota. El error se da en muestras. El conjunto de datos utilizados para entrenamiento es de 16429 trazas y validación es de 4108 trazas, se determinan los aciertos de las redes y errores entre una a cinco muestras y superiores a 5 muestras.

Las implementaciones de las dos estrategias (DNN y CNN) parten de condiciones comunes como: igual rango de SNR añadido, mismo número de trazas en entrenamiento y validación y la aplicación de una transformación para entrenar a la red bajo otro dominio. La implementación completa de cada una de las estrategias cuentan con parámetros adicionales que son necesarios para su aplicación, como lo son la definición correcta de las diferentes transformadas propuestas por cada uno de los artículos. La selección de la red convolucional se debe al desempeño de los resultados de la CNN frente a la red DNN. De acuerdo con la función de pérdida, para el caso de la DNN mostrada en la Figura 15, visualizamos que se presenta *overfitting*. Este comportamiento usualmente ocurre porque nuestro modelo considera como datos válidos únicamente aquellos que sean similares a los datos utilizados en el conjunto de entrenamiento, de tal forma que no puede

distinguir otras características de fiabilidad en entradas a la red, si estas varían de los rangos con los que fue entrenada. Otro de los factores dentro de la selección de la arquitectura CNN corresponde a la cantidad de transformaciones sugeridas por los artículos. Para la DNN, Mężyk y Malinowski (2019) propone 11 diferentes transformaciones las cuales requieren de una cantidad considerable de tiempo para su ejecución e implementación, dentro del Capítulo 3 la sección 3.1 es un trabajo inacabado, ya que sólo se llevó a cabo una de las 11 transformaciones utilizadas por los autores. La ejecución e implementación de cada una de las transformaciones implica la selección y adecuación de parámetros necesarios dentro de la ecuación que describe el comportamiento de la transformación, que no son genéricos dentro de ellas para llevarlas a cabo, situación que no nos beneficia abordar totalmente puesto cada transformación tiene diferentes expresiones de acuerdo al tipo de estudio requerido. Caso contrario en la implementación de la CNN en donde, el uso de la transformada Wavelet escalada a un nivel específico, implica ya contar con las transformadas predecesoras a la calculada empleando herramientas ya estipuladas como la *cwt en Python*. Si bien, el trabajo computacional de la Wavelet es mayor, comparada con realizar 11 transformaciones diferentes a una misma traza, como es propuesto en la DNN, es más sencilla de realizarse. Teniendo en cuenta estos criterios y de acuerdo a los resultados obtenidos, se ha decidido, dentro de la detección de primeros arribos propuesta en este proyecto, utilizar la topología de red convolucional (CNN) y a partir de lo ya mostrado, variar parámetros y criterios para la entrega del prototipo diseñado.

La red CNN cuenta con la ventaja de mezclar diferentes tipos de capas (utiliza capas profundas y capas convolucionales) que le permiten detectar características con altos niveles de abs-

tracción a lo largo de cada capa, y una vez aprendida, la convolucional la reconoce en cualquier punto de la red. Frente a la red DNN que sólo usa capas densas que aprenden los patrones por capa y que debe aprender nuevamente las características si se presenta en otro punto de la red. El uso de la CNN beneficia el entrenamiento de la red, ya que permite la adquisición de más características en su paso por cada capa. La diferencia fundamental entre una capa densamente conectada y una convolucional corresponde al reconocimiento de patrones, para el caso de la DNN la red aprende tomando el conjunto total de datos de entrada, a diferencia de la CNN que identifica las características analizando vecindades a la muestra de interés, detectando así patrones por pequeñas ventanas que recorren los datos de entrada (Torres, 2018).

4. Red solución para la detección de primeros arribos

La arquitectura de red propuesta como solución para la detección de primeros arribos corresponde a una red convolucional basada en las implementaciones expuestas en el Capítulo 3. Para el diseño de la red solución partimos de la arquitectura planteada en la Figura 21, a partir de esta arquitectura, se varían las configuraciones de parámetros e hiperparámetros utilizados en la CNN y se define la estructura de capas necesarias para optimizar el proceso de detección. De las implementaciones realizadas durante las pruebas, se determina que los resultados de las redes en el Capítulo 3 pueden tener mejoras al unirse las redes CNN con la DNN. Esta unión permite crear la red solución, la cual está compuesta por una capa convolucional 1D, una capa *maxpooling* 1D, una capa *flatten* y la distribución de las capas ocultas de la DNN. Esta organización es definida de forma completa a partir de nuevas pruebas en las cuales se valida la selección de parámetros e hiperparámetros, esta validación se realiza bajo múltiples iteraciones en las cuales, son asignados el número de épocas de entrenamiento, la función de pérdida, función de activación, optimizador, número de entradas, entre otros.

Las pruebas de las diferentes configuraciones dentro del proceso del diseño de la red solución, se encuentran disponibles para su revisión en el repositorio de GitLab del Grupo de Investigación en Conectividad y Procesamiento de Señales (CPS), en donde se encuentra el resultado de

implementar todas las pruebas y el diseño de la red catalogada como solución.

4.1. Experimentos y comparación de resultados

Para el diseño de la red solución las trazas utilizadas corresponden al conjunto de datos descrito en la Tabla 4, los cuales conservan una SNR de 6 a 10 dB con distribución normal a lo largo de la traza. El diseño de la red solución parte de la realización secuencial de tres grandes conjuntos de experimentos. Los experimentos parten de las siguientes hipótesis:

- I. Los parámetros e hiperparámetros utilizados en la configuración de la red CNN descrita en el Capítulo 3 deben ser ajustados para mejorar los resultados en implementaciones para trazas con 1251 muestras.
- II. La arquitectura de la red neuronal puede entregar mejores resultados si se disminuyen las capas que la componen, es decir, no contar con la réplica de capas convolucionales ni *maxpooling* presentadas por la estrategia; para comprobar la robustez del modelo y la contribución de las capas replicadas en la detección de los primeros arribos.
- III. Una vez comprobado el desempeño de la red relacionado con el número de capas dentro de la arquitectura, se propone la modificación nuevamente de parámetros e hiperparámetros para lograr obtener de forma eficiente y con un alto nivel de acierto el primer arribo.

El primero conjunto de experimentos consiste en verificar si la estructura de la red CNN propuesta en el Capítulo 3 en conjunto con las DNN entregan mejores resultados a partir del cambio

de parámetros. De la mejor configuración dentro de este primer conjunto de experimentos, se procede a realizar el análisis de la estructura variando el número de capas y la secuencia que se realiza para la construcción de la red solución. Y finalmente, con los resultados de este segundo conjunto de experimentos, se varían nuevamente los parámetros e hiperparámetros para seleccionar los más convenientes dentro de la implementación de la red y posteriormente ser entregados como arquitectura final.

4.1.1. Experimentos red convolucional

Como se expresa al final del Capítulo 3, el diseño de la red solución parte de la configuración convolucional representada por la Figura 21. A partir de la CNN se realiza la implementación de experimentos con la misma estrategia de la Sección 3.2, en la que se hace uso de un filtro pasa bajas y de 31 transformaciones Wavelet. Cada experimento consiste en la variación de parámetros e hiperparámetros con el propósito de asignarlos de forma adecuada en la implementación de la red solución. Para esto, se genera una serie de 10 experimentos expuestos en la Tabla 7, en donde cada experimento es entrenado a partir de una etiqueta la cual consiste en una máscara de unos y ceros en cada muestra (Figura 37). En la Tabla 7 son expuestas las diferentes variaciones realizadas a la CNN original y cada prueba se encuentra organizada en una carpeta de GitLab por medio de un hipervínculo en el nombre del experimento ubicado en la columna "Modelos de Red".

Durante el entrenamiento de los experimentos enunciados en la Tabla 7, se evidencia una mejora progresiva de los resultados a la salida de cada red.

Tabla 7

Diseño de experimentos para la selección de parámetros de la CNN

Diseño de experimentos para la selección de parámetros de la CNN	
Modelos de Red	Experimento
Función de pérdida error cuadrático medio	Original
Función de pérdida error absoluto medio	E2
Dropout del 10 % en la primera capa de <i>maxpooling</i>	E3
Red con la mitad de capas ocultas	E4
Dropout del 10 % en capas convolucionales y capas densas	E5
Dropout del 10 % en capas densas	E6
Función de activación ReLu en todas las capas	E7
Función de activación LeakyRelu en todas las capas	E8
Función de activación LeakyRelu en capas densas	E9
Función de activación ReLu en capas densas	E10

Nota. Variaciones basadas en la red convolucional ilustrada en la Figura 21. En el experimento E4, la red consta de la mitad de capas ocultas (secuencia de capas ocultas de la red resultante: capa convolucional, max pooling y capa densa).

La obtención de salidas cada vez más similares al etiquetado, permite realizar la selección de la configuración de red para el experimento con mejor desempeño, mediante la comparación de los resultados de la función de pérdida y los errores de las predicciones entregadas por cada red. En la Tabla 8 se encuentran los valores obtenidos de las métricas para los datos tanto de entrenamiento como de validación en cada experimento. En la Tabla 8 las métricas de los experimentos siete y diez no son calculadas debido a que sus salidas arrojan resultados atípicos, son variaciones periódicas que no siguen la etiqueta de entrenamiento y que dificultan el uso de la primera diferencia como método de identificación del primer arribo. Para poder evaluar los resultados de estos experimentos es necesario variar el criterio de selección del FB que se ajuste de forma adecuada a

Tabla 8

Errores con datos de entrenamiento y validación, red CNN

Datos	Experimento	Media	Desviación Estándar
Entrenamiento	Original	19	40
	E2	53	90
	E3	38	82
	E4	73	103
	E5	41	75
	E6	46	71
	E8	65	85
	E9	69	103
	Validación	Original	20
E2		54	93
E3		38	42
E4		70	100
E5		41	74
E6		45	70
E8		67	88
E9		70	102

Nota. Las métricas fueron calculadas en términos de muestras. La media y desviación estándar se obtiene con base en el cálculo de la primera diferencia en la salida de cada red, dando como resultado final la predicción del FB en la muestra de máxima diferencia.

las características de sus correspondientes salidas, la variación del criterio de selección amplía el número de comparaciones requeridas para el análisis e implica aplicar el criterio a más experimentos para poder evaluar los resultados en las mismas condiciones, siendo requerido mayor tiempo para el diseño de la red solución.

En la Tabla 8 se evidencia que con datos de validación, la media y la desviación estándar

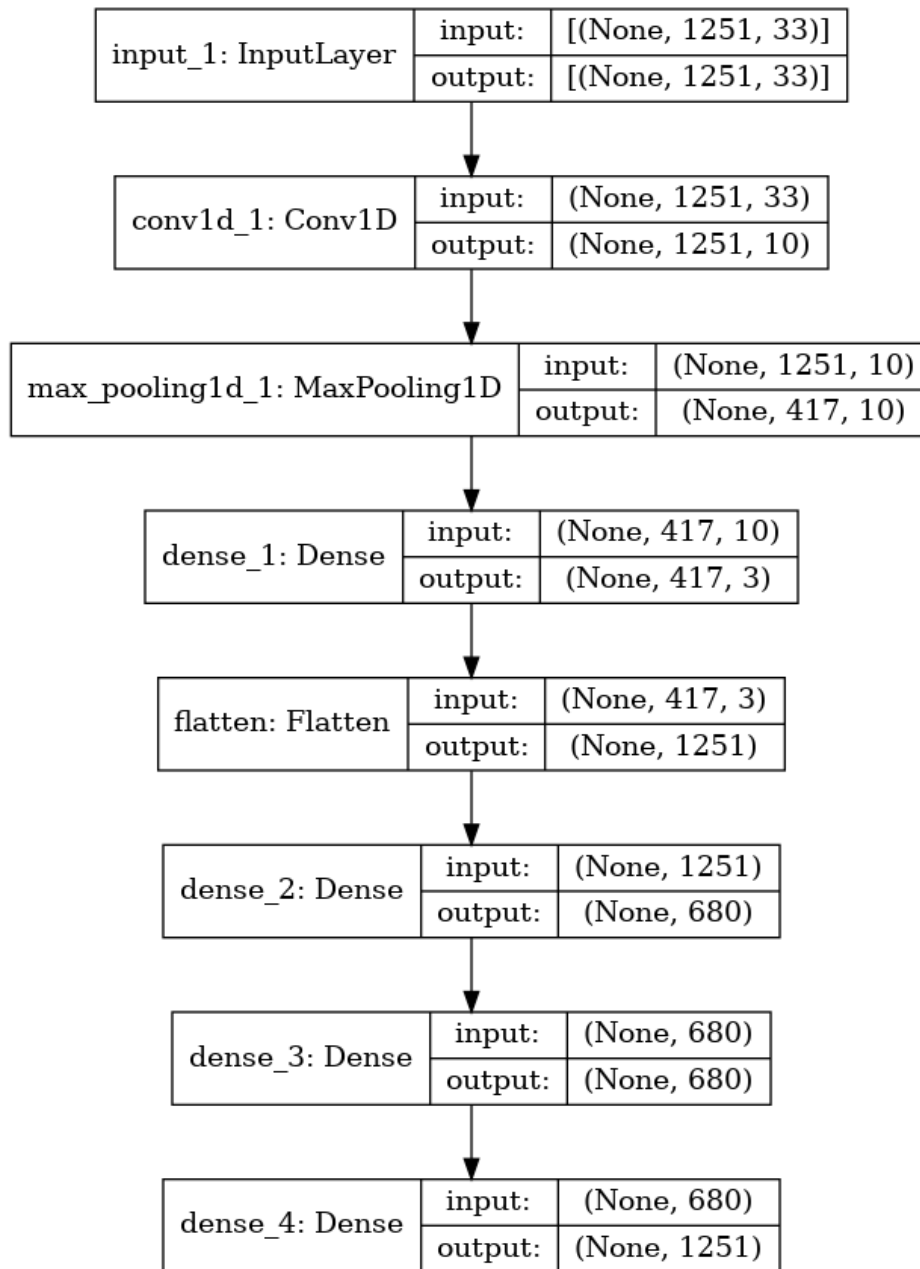
mantienen su comportamiento respecto a la predicción realizada con datos de entrenamiento. Al hacer una comparativa sobre cada uno de los experimentos planteados, se corrobora que el mejor experimento corresponde a la estructura de la CNN original expuesta en la Figura 21.

Con lo mencionado anteriormente, se genera una red con características del experimento original de la Tabla 8 para el conjunto de trazas originales con 1251 muestras de longitud. Se plantea optimizar la clasificación de la CNN reestructurando su arquitectura, manteniendo las cuatro primeras capas ocultas de la CNN del modelo (Figura 21) y reemplazando las últimas cinco restantes por las tres capas ocultas de la DNN (Figura 12), como se ilustra en la Figura 31. Para contrastar esta hipótesis se plantean tres experimentos descritos en la Tabla 9, manteniendo las 33 señales de entrada descritas en el Capítulo 3.

En la Tabla 9 se registran tres variaciones de estructura para contrastar la mejor opción entre usar sólo las tres primeras capas ocultas de la arquitectura expuesta por la Figura 21, mantener las cuatro primeras capas ocultas de la CNN y reemplazar las últimas cinco restantes por las tres capas ocultas de la DNN, o agregar *batch normalization* (en cuanto a predicción). Dicha tabla muestra que los experimentos que sobresalen por su bajo error con datos que ya conocía la red y los que no ante las tres variantes, son los experimentos 2 y 3 con errores muy cercanos.

Figura 31.

Arquitectura red solución: unión entre CNN y DNN



Nota. La arquitectura unión consta de la mitad de capas ocultas de la red CNN (Figura 21) y se realiza la unión entre la CNN y DNN.

Tabla 9

Métricas de resultados variando estructuralmente la unión de CNN con DNN

Datos	Experimento	Media	Desviación Estándar
Entrenamiento	1	137	199
	2	11	9
	3	6	7
Validación	1	133	189
	2	11	9
	3	7	7

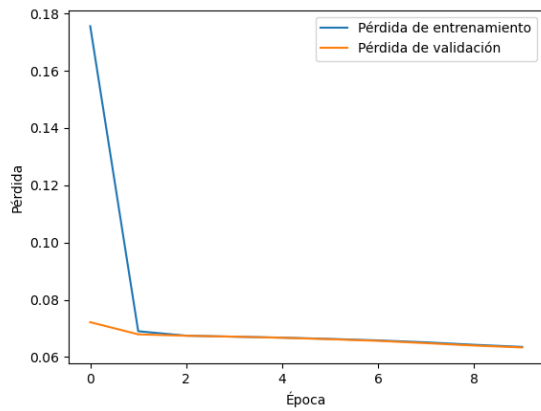
Nota. Errores de predicción en tres variantes estructurales de la CNN. Experimentos con 10 épocas de entrenamiento: Experimento 1) tres capas ocultas constituidas por capa convolucional 1D, maxpooling 1D y capa densa, experimento 2) unión de la CNN con la DNN (Figura 31) y experimento 3) se agrega Batch normalization después de cada capa oculta del experimento 2.

En la Figura 32 se muestran las funciones de pérdida del entrenamiento de los experimentos de la Tabla 9, la cual en el experimento 2 (b) se logra evidenciar en la última época que puede seguir aprendiendo a comparación del experimento 3 (c) en el que se logra ver que el aprendizaje de la red se tiende a estabilizar. En consecuencia, por el rendimiento y la tendencia del bajo error de entrenamiento del experimento 2, se toma la decisión de seguir estudiando esta arquitectura (Figura 31). Los parámetros e hiperparámetros que conforman la arquitectura de los tres experimentos corresponden al resultado Original realizado en las variaciones mostradas en la Tabla 8. Cabe aclarar, que éstas no necesariamente corresponden a la selección más adecuada, ya que son arquitecturas diferentes respecto a los experimentos ya planteados, y por tanto, se debe estudiar y contrastar cuáles son los parámetros e hiperparámetros más adecuados para el diseño de la red.

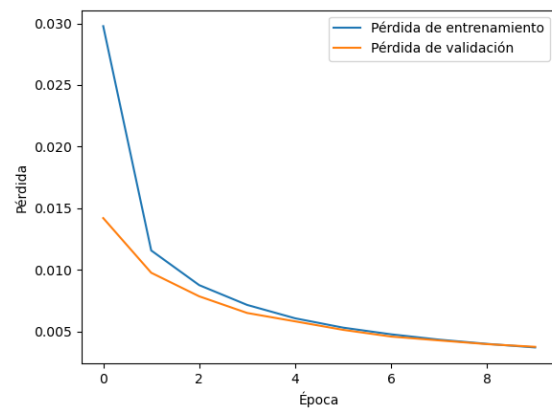
Las gráficas de las funciones de pérdida fueron determinadas por medido de 10 épocas de

Figura 32.

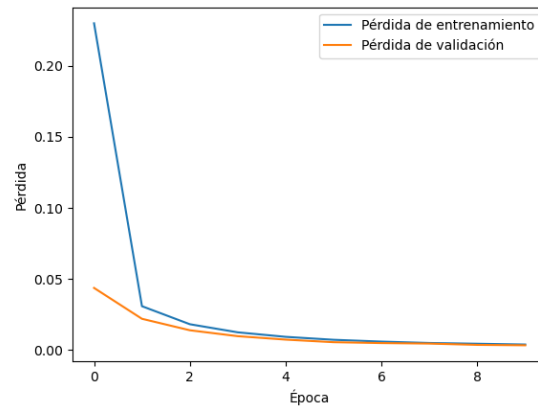
Funciones de pérdidas en entrenamiento y validación de las variaciones de estructura en la unión de CNN con DNN



(a) Función de pérdida y validación, experimento 1.



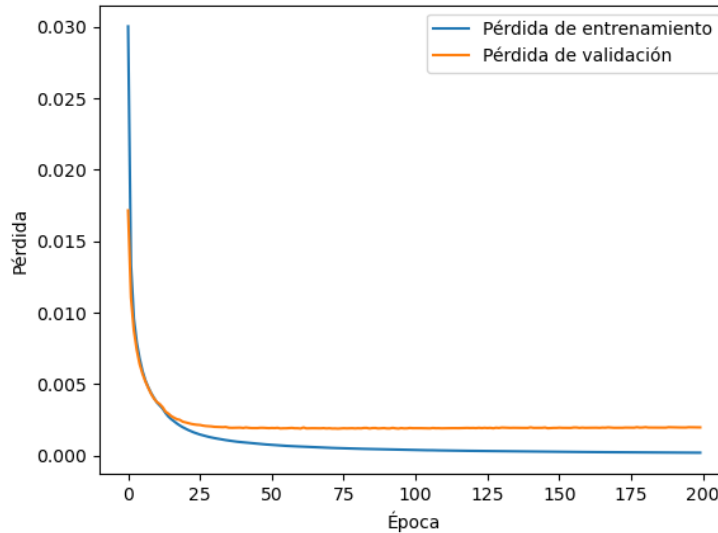
(b) Función de pérdida y validación, experimento 2.



(c) Función de pérdida y validación, experimento 3.

entrenamiento con el objetivo de visualizar su progresión de forma inicial. En consecuencia con la Tabla 9 y la Figura 32, se continua el estudio del experimento 2, expandiendo el número de épocas de 10 a 200 en el entrenamiento como muestra la Figura 33, con el propósito de ver en qué época finaliza el transitorio de las curvas de pérdida y validación del entrenamiento y determinar hasta

Figura 33. Función de pérdida de la unión de las redes CNN con DNN



Nota. Función de pérdida para un entrenamiento con 200 épocas.

qué época es conveniente entrenar la red antes de que exista un sobreajuste. Del resultado obtenido para 200 épocas, en la Figura 33 se logra visualizar que a partir de la época 25 el aprendizaje tiende a ser estable, sin embargo, con el aumento de las épocas la diferencia entre la curva de pérdida y validación del entrenamiento tiende a ser constante (existe una separación entre las curvas) lo cual puede llevar a disminuir la generalización del modelo y su vez a una predisposición de sobreajuste. Por lo tanto, el modelo es entrenado con 25 épocas.

4.1.2. Variación de parámetros e hiperparámetros de la CNN unida con la DNN

Habiendo escogido la mejor estructura bajo los múltiples experimentos realizados, se procede a contrastar y seleccionar los parámetros e hiperparámetros que mejor se ajustan a la estructu-

ra de la red solución (Figura 31). Para ello se hace un total de 38 experimentos (20 correspondientes a la “Serie a” y 18 a la “Serie b”) registrados en la Tabla 11. La “Serie a” corresponde a la variación de parámetros de la red bajo la arquitectura del Experimento b en la Tabla 9 en la que se define la función de activación y a partir de esta configuración de parámetros se realizan 18 nuevos experimentos denominados "Serie b". Con este nuevo conjunto de pruebas se define el número de neuronas que se necesitan para la red solución, además de justificar los parámetros escogidos para su implementación.

Los experimentos de la "Serie a" se efectúan respecto a la configuración anteriormente descrita por la Figura 31. En la Tabla 11 se visualizan los resultados de las métricas para cada experimento. En los experimentos 1 y 2 se varía la función de pérdida de (1) Error cuadrático medio a (2) Error absoluto medio. En los experimentos 3 al 5 se aplica *dropout* del 10% en diferentes capas, se aplica (3) *dropout* sólo en la capa *maxpooling*, (4) *dropout* en la capa *maxpooling* y las capas densas, (5) *dropout* sólo después de las capas densas. Para las siguientes pruebas, se varía la función de activación de las tres primeras capas ocultas, (6) se aplica la función de activación Relu y en el experimento (7) se aplica la función de activación LeakyReLU con valores por defecto (α de 30%). En los experimentos (8) al (13), en las últimas capas densas se cambia la cantidad de neuronas: 200, 500, 700, 900, 1300, 2600 en cada experimento. En los experimentos (14) al (18) se modifica el *batch size* en los valores 50, 75, 200, 470 y 600 respectivamente. Por último, en 19 y 20 se varía la función de activación, para (19) en la primera capa densa se cambia de “tangente hiperbólica” a *sigmoid*, y en (20) las dos últimas capas ocultas se varía de *softsign* a *sigmoid*.

En los experimentos (1) y (2), se destaca una menor media de error y baja desviación estándar del experimento (1), tanto para datos de entrenamiento como de validación, al cambiar la función de pérdida. En los experimentos (3) al (5) aplicar el *dropout* sólo después de las capas densas, entrega en los resultados finales un aumento en la media y desviación estándar respecto al experimento (1). La actualización de los pesos sin considerar neuronas que han sido abandonas contribuyen a la sensibilidad del modelo y mejoran la generalización de los datos para evitar el sobreajuste, sin embargo, el aumento en las métricas puede disminuir y mejorar los resultados si se aplica *dropout* a la capa de entrada. El cambio de la función de activación de las tres primeras capas ocultas, (6) se aplica la función de activación *Relu* y en el experimento (7) se aplica la función de activación *LeakyReLU* con valores por defecto (α de 30%); respecto al experimento (1), en (6) y (7) se logra mantener la media de los datos de entrenamiento y disminuir en los datos de validación, adicionalmente la desviación estándar se logra disminuir.

Para los experimentos (8) al (13), permiten visualizar que respecto al experimento (6) y (7) en las distintas variaciones de neuronas se refleja un aumento en la media con datos de validación y a su vez en la desviación estándar tanto con datos de entrenamiento como de validación. Para los experimentos (14) al (18) se varía el *batch size*, los resultados indican que a medida que se aumenta el *batch size*, tanto para datos de validación como de entrenamiento, se presenta una tendencia al aumento de la media y desviación estándar que son obtenidas a partir de los resultados que entrega la red posterior a la salida y la primera diferencia realizada. Finalmente, en los experimentos (19) y (20) se cambia la función de activación, respecto al experimento (6) y (7), la media y desviación

estándar aumentan tanto con datos de entrenamiento como de validación.

Tabla 10

Análisis de sensibilidad en la identificación de parámetros e hiperparámetros

Parámetro	Observaciones
Función de pérdida	El MSE permite enfatizar los grandes errores presentes en la predicción de primeros arribos, siendo notorias las variaciones de los resultados frente al entrenamiento e identificando así, condiciones especiales en trazas en las que se presentan dichas variaciones.
<i>Dropout</i>	La aplicación de <i>Dropout</i> permite contar con un entrenamiento más complejo y robusto identificando características prevalentes contribuyendo en la convergencia del modelo. Observamos que su aplicación debe realizarse a las capas ocultas de la red y a las entradas de la misma para mejorar resultados
<i>Batch size</i>	El aumento del tamaño del batch a valores más grandes implica el uso de más recursos computacionales y requiere de más tiempo para el entrenamiento.
Función de Activación	Este parámetro afecta la linealidad a la cual son sometidos los datos incide directamente en el desempeño de la red afectando en las métricas de acuerdo a la función utilizada y su ubicación en las capas de la red.

Nota. Observaciones encontradas durante la implementación de los diferentes experimentos para la selección de parámetros dentro del análisis de sensibilidad.

En la Figura 34, se puede evidenciar que los experimentos con menor error en predicción para datos de entrenamiento y validación corresponde a los experimentos 6, 7 y 13. Para los experimentos 6 y 7 se logra apreciar una tendencia a menor media con trazas de entrenamiento y menor

Tabla 11

Errores variando parámetros e hiperparámetros de la unión de CNN con DNN

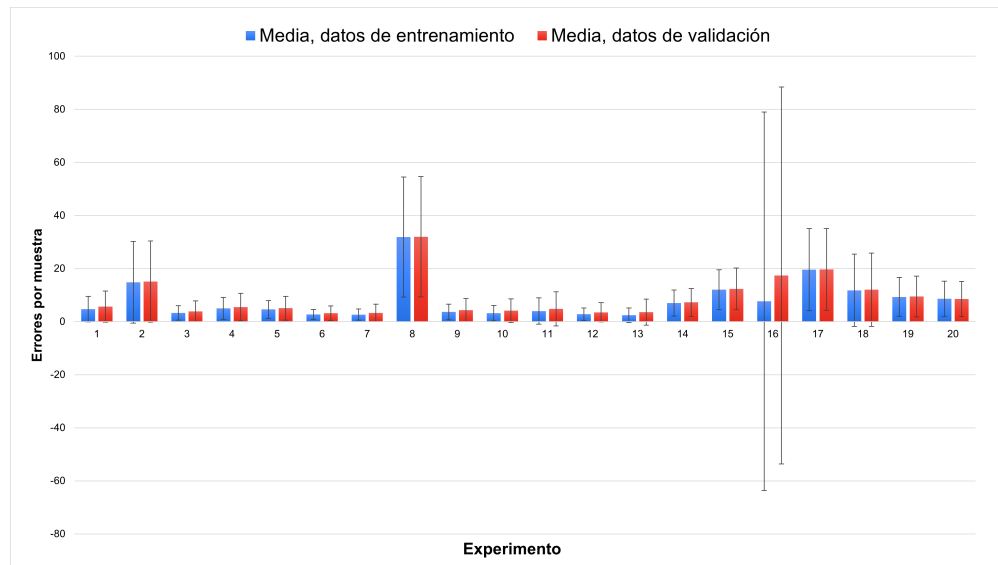
Experimento	a				b			
	Media		Desviación estándar		Media		Desviación estándar	
	D.E.	D.V.	D.E.	D.V.	D.E.	D.V.	D.E.	D.V.
1	5	6	5	6	3	3	2	3
2	15	15	15	15	14	14	13	13
3	3	4	3	4	3	4	2	3
4	5	5	4	5	6	6	5	7
5	5	5	3	4	4	5	3	4
6	3	3	2	3	/	/	/	/
7	3	3	2	3	/	/	/	/
8	32	32	23	23	5	5	3	4
9	4	4	3	4	3	4	2	4
10	3	4	3	4	3	3	2	3
11	4	5	5	6	2	3	2	3
12	3	4	2	4	2	3	2	4
13	2	4	3	5	3	5	5	9
14	7	7	5	5	2	3	3	4
15	12	12	8	8	2	3	2	3
16	8	17	71	71	3	3	2	3
17	20	20	16	15	5	6	4	4
18	12	12	14	14	7	8	5	5
19	9	9	7	8	11	11	8	8
20	9	8	7	7	14	14	10	10

Nota. Las siglas D.E. corresponden a datos de entrenamiento y D.V a datos de validación. En cada serie de experimentos se calculan dos medidas de error (Media y Desviación estándar) con dos conjuntos de datos diferentes (trazas de Entrenamiento y Validación). La “serie a” consta de 20 experimentos con la arquitectura escogida en el Experimento 2. La “serie b” cuenta con 18 experimentos los cuales surgen de la configuración con mejor desempeño en la “serie a”. Cada iteración permite corroborar el parámetro e hiperparámetro más adecuado para la red solución.

desviación estándar con trazas de validación, lo cual recalca la precisión con trazas no conocidas manteniendo su exactitud de predicción. En el experimento 13, al aumentar al doble la cantidad de neuronas en las dos últimas capas ocultas, la media con datos de entrenamiento se mantiene, mientras que con datos de validación aumenta, y a su vez, tanto para datos de entrenamiento como de validación la desviación estándar aumenta. De acuerdo con los datos encontrados mediante la evaluación de las redes con datos de validación, se estableció que la mejor arquitectura consiste en el experimento 7, debido a que el modelo logra una mejor generalización durante el entrenamiento frente al experimento 6 y comparado con el experimento 13 en el que el aumento de la desviación disminuye la exactitud y precisión en la tarea de detección de FB.

Figura 34.

Histogramas con métricas de error en predicción de la “Serie a” de la Tabla 11.



Nota. La desviación estándar de predicciones son proporcionales entre datos de entrenamiento y validación y está relacionada respecto a la media, es representada por las líneas de error verticales para cada medida en los experimentos.

Por consiguiente se logra determinar que la configuración con mejor exactitud y precisión

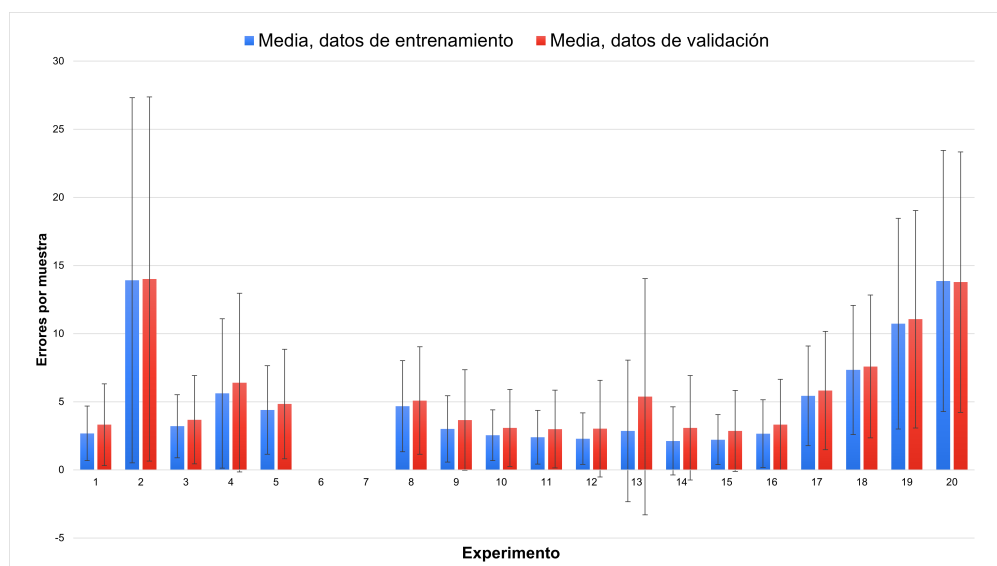
respecto a las variantes expuestas en “Serie a” de la Tabla 11 es la expuesta por el experimento 7. A partir de esta configuración, se genera una serie de 18 experimentos semejantes a los descritos en “Serie a” registrados bajo el nombre de pruebas “Serie b”, las cuales mantienen los mismos criterios y variaciones en los experimentos (1) al (13), para los experimentos del (14) al (20) se varía el *batch size* por 5, 10, 15, 50, 75, 200 y 470, respectivamente. De los experimentos realizados para la “Serie b” se omiten las pruebas correspondientes a la variación de la función de activación para las tres primeras capas ocultas que fueron desarrolladas en los experimentos 6 y 7 de la “Serie a”.

En la Figura 35 se evidencia que el comportamiento en los experimentos de la “Serie b” para las pruebas (1) a la (13) se logra mantener el comportamiento analizado en los resultados del conjunto de “Serie a”, y los experimentos del (14) al (20) se logra visualizar que existe un aumento del error sobre las predicciones realizadas tanto en datos de entrenamiento como de validación. La ejecución de los “Serie b” permite destacar la configuración de los experimentos 10 (respecto al experimento 1 varía a de 680 a 700 neuronas en las últimas dos capas densas ocultas) y 16 (respecto al experimento 1 varía el tamaño del *batch* de 20 a 15).

Los resultados obtenidos son buenos, sin embargo, no entregan aportes significativos frente al experimento 1 considerado como mejor opción, ya que las distintas variaciones de parámetros alcanzan valores de media y desviación estándar cercanos al experimento 1, pero requieren de un mayor costo de implementación.

Figura 35.

Histogramas con métricas de error en predicción de los experimentos en la “Serie b” de la Tabla 11

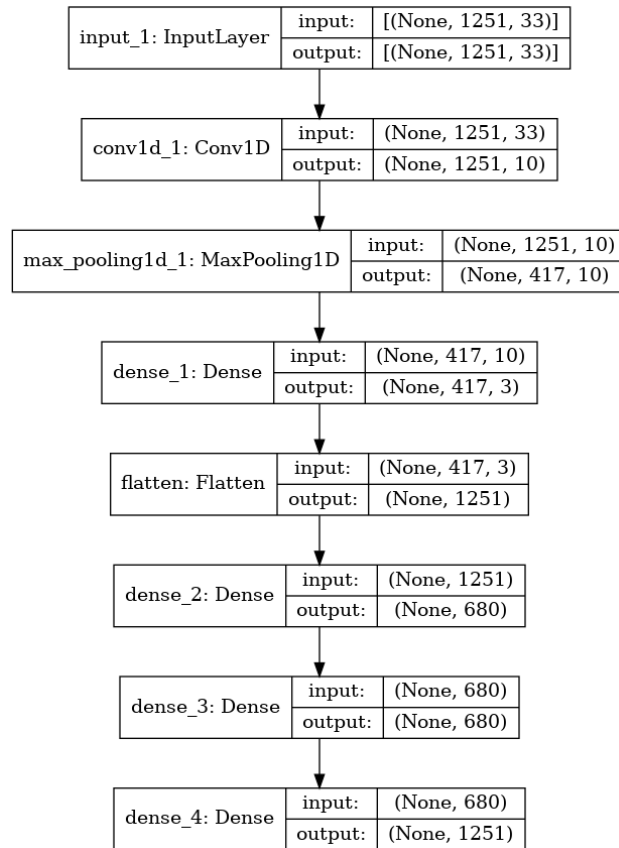


Nota. La desviación estándar de predicciones son proporcionales entre datos de entrenamiento y validación y está relacionada respecto a la media, es representada por las líneas de error verticales para cada medida en los experimentos.

4.2. Red solución para predecir el primer arribo

De los experimentos anteriormente realizados, se escoge la red solución (conformada por la unión de la CNN con la DNN) descrita por la Figura 36. El conjunto de trazas descritas en la Tabla 1 cuyos FB se encuentren desde la muestra 50 en adelante, con ruido y normalizadas, es utilizado para el entrenamiento y validación de la red solución. Para este proceso, se mantiene el uso de 31 transformaciones Wavelet, filtro pasa baja, SNR y normalización de la traza para señales con 1251 muestras.

Figura 36. Arquitectura de la red solución

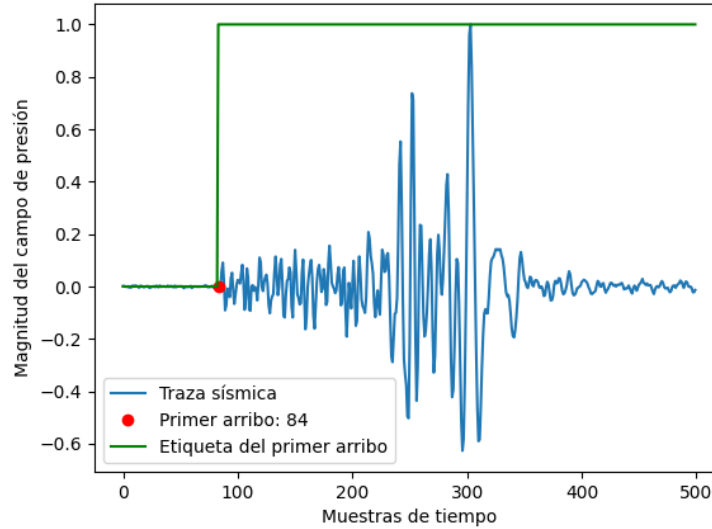


Nota. La red cuenta con una capa convolucional, seguida de una *maxpooling*, en conjunto de una capa densa con la suma de la red DNN definida en la sección 3.1. Este procedimiento se realiza con el objetivo de afianzar el modelo de aprendizaje siendo una red robusta ante el ruido para la detección del FB.

Durante el entrenamiento, se entregan en conjunto el vector de amplitudes con 1251 muestras por traza y una etiqueta del primer arribo con unos y ceros de 1251 muestras, a la red convolucional, Figura 37. Para el etiquetado del FB las muestras anteriores al primer arribo tienen un valor de cero y para las muestras posteriores, incluyendo el primer arribo, tiene un valor de uno.

Figura 37.

Etiqueta de entrenamiento en la red solución

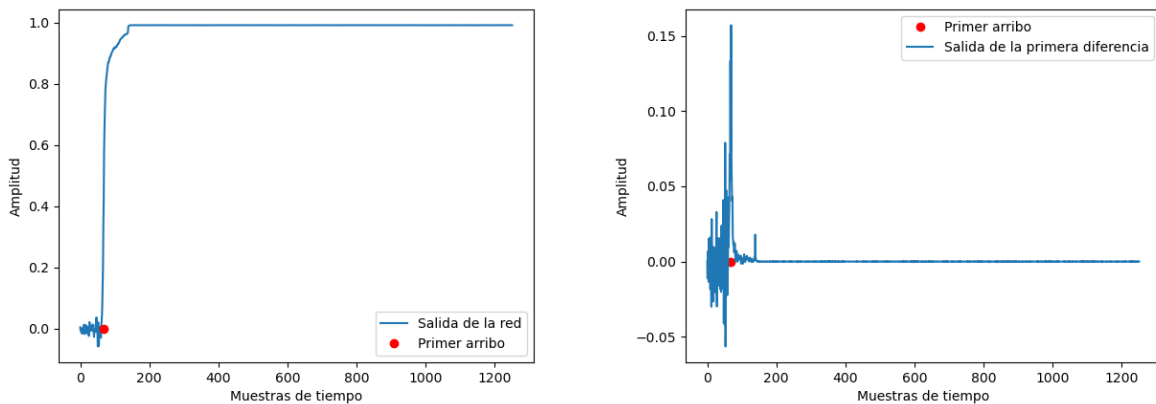


Nota. El etiquetado (máscara del primer arribo representado en verde) del entrenamiento consiste en una función escalón unitario con una longitud de 1251 valores asignando un cero a las muestras antes del FB y uno a las muestras siguientes incluyendo el primer arribo.

Partiendo de la salida ideal para el escalón unitario durante su aprendizaje, la red solución aprende la máscara que ha sido empleada e intenta adaptar los pesos en cada época para obtener a la salida un escalón lo más parecido al utilizado durante el entrenamiento. La salida de la red consta de una sola capa, la cual entrega un vector de longitud 1251 en la que cada posición representa un valor de probabilidad que indica el nivel de semejanza del valor obtenido frente al esperado. Teniendo en cuenta que la máscara que se usa para realizar el entrenamiento es un escalón unitario en la posición del FB, se aplica la primera diferencia a la salida de la red con el fin de identificar la posición más probable del FB como la mayor primera diferencia encontrada. La posición en la

que la diferencia tenga un valor máximo corresponde a la detección del primer arribo. La máxima diferencia es seleccionada como la primera de resultados de izquierda a derecha, es decir, en caso de que la máxima diferencia se obtenga en más de una de las muestras, se toma la primera y el resultado se asocia a la muestra en la que se detecta el máximo por primera vez. Esto debido a que los máximos encontrados posteriores se catalogan como falsos primeros arribos.

Figura 38. Representación gráfica de la salida y predicción de la red solución



(a) Salida de la red solución para el procesamiento de la traza n° 22.

(b) Resultado de primera diferencia a la salida de la red solución para la traza n°22.

Para la elección de los parámetros se tuvieron en cuenta los siguientes criterios: disminución en la función de pérdida en el modelo de la red solución, resultado de las métricas y visualización de primeros arribos generados. Los parámetros utilizados para la implementación de la red solución, fueron seleccionados a partir de las diversas iteraciones y experimentos realizados en la sección 4.1 del Capítulo 4 y son mostrados en la Tabla 12, 13 y 14.

Tabla 12

Definición de hiperparámetros para la arquitectura de solución propia

Función de Activación	Optimizador	Métrica	Batch size	Capas	Función de pérdida
LeakyRelu	Adamax	Binary accuracy	20	5	MSE [mean squared error]

Nota. Los hiperparámetros fueron seleccionados a partir del resultado de una serie de experimentos expuestos en la sección 4.1. El número de capas corresponden a 1 capa convolucional, 1 maxpooling, 2 capas densamente conectadas que corresponden a la DNN implementada en el Capítulo 3 y 1 reshape. La función de activación LeakyRelu se aplica a las tres primeras capas ocultas de la red.

Tabla 13

Selección de parámetros para el diseño de la red solución.

Tamaño del Kernel	Número de Kernel	Padding	Stride	Función de Activación
20	10	<i>same = 1251</i>	1	Softsign

Nota. Los parámetros para la red solución fueron seleccionados basados en los resultados de la serie de experimentos descritos en la Sección 4.1.

Tabla 14

Definición de hiperparámetros para la arquitectura DNN usada en red solución

Función de Activación	Óptimizador	Inicializador	Batch size	Capas ocultas	Número de neuronas	Función de pérdida
Softsign	Adamáx	Glorot uniform	20	2	680	MSE

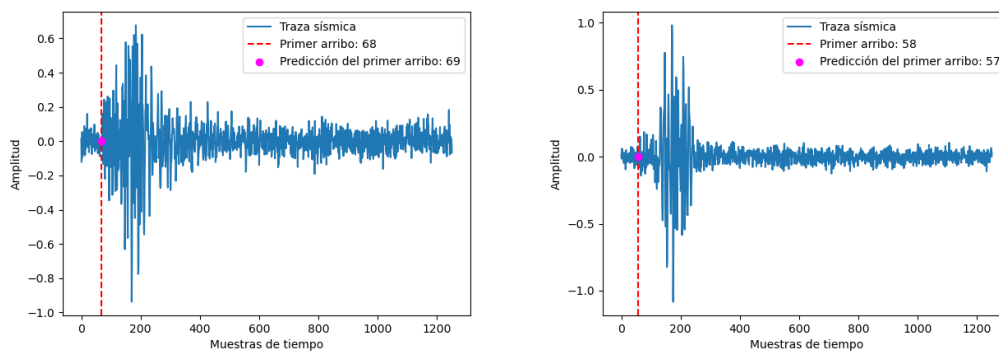
Nota. Los hiperparámetros fueron seleccionados a partir del resultado de la evaluación de desempeño para cada uno de los experimentos realizados.

La red solución cuenta con un tiempo de entrenamiento de 3 minutos con 66 segundos y un tiempo de predicción para 4108 datos de 24,07 segundos. En este tiempo, se utilizan entre un 17 a 28 % de recursos de GPU y un 100 al 458 % de recursos en CPU.

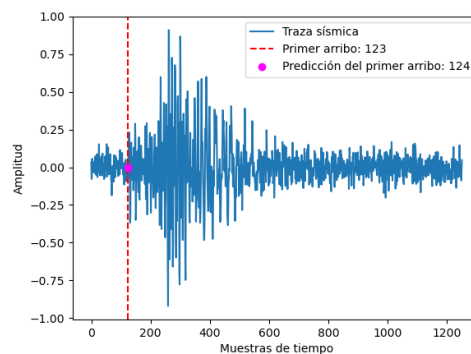
Una vez realizada la primera diferencia, se selecciona el máximo valor entregado por la resta de la muestra actual menos la anterior (Figura 38). El valor máximo de la diferencia corresponde al primer arribo ya que, al ser etiquetado con un escalón unitario, el resultado de la primera diferencia nos entregará un pulso justo en la muestra FB. Finalmente, el valor que se predijo por la implementación de la red, es visualizado por medio de la Figura 39.

Figura 39.

Estimación del valor del primer arribo con la red solución



(a) Predicción del primer arribo de la traza 22. (b) Predicción del primer arribo de la traza 2022.

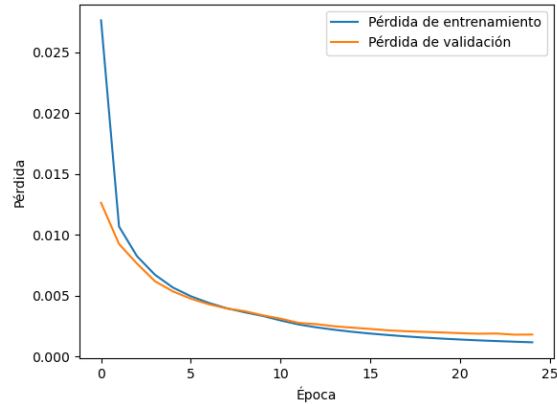


(c) Predicción del primer arribo de la traza 4015.

Nota. Predicción del primer arribo en tres trazas escogidas de forma aleatoria sobre el conjunto de 4108 trazas de validación. El primer arribo visualmente se sobrepone en la muestra correspondiente al primer arribo de la traza el bajo error (una muestra de error) de detección con la red solución.

Figura 40.

Función de pérdida de entrenamiento y validación en red solución.



Nota. El aprendizaje de la red se realiza para 25 épocas. La curva de pérdida en el conjunto de validación inicia en un valor cercano a 0.013 frente a un valor de entrenamiento de 0.027.

A partir de la realización del entrenamiento de la red solución, se obtienen las gráficas correspondientes a la función de pérdida para cada conjunto de datos: entrenamiento y validación, presentadas en la Figura 40. Con los resultados obtenidos en la predicción para el conjunto de 4108 trazas, se calcularon diferentes métricas para evaluar el desempeño de la red. Los cálculos realizados son mostrados en la Tabla 15.

Tabla 15

Métricas de resultados para red solución con datos de entrenamiento y validación

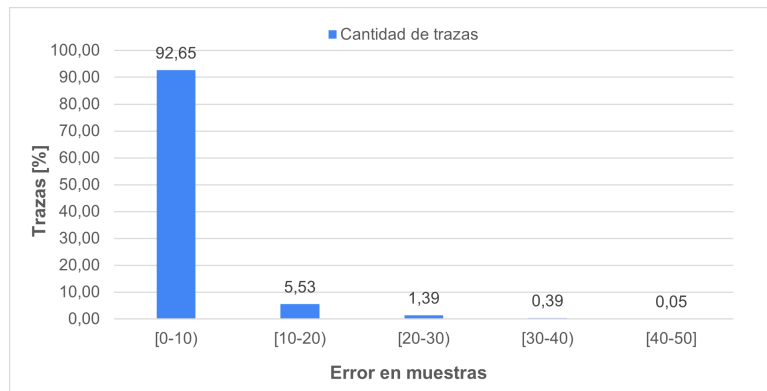
Datos	Media	Desviación estándar
Entrenamiento	3	4
Validación	4	5

Nota. Las métricas fueron calculadas en términos de muestras. Para el caso de la media, se determinó como el error absoluto entre el resultado del picado verdadero de los datos frente al picado predicho por la red.

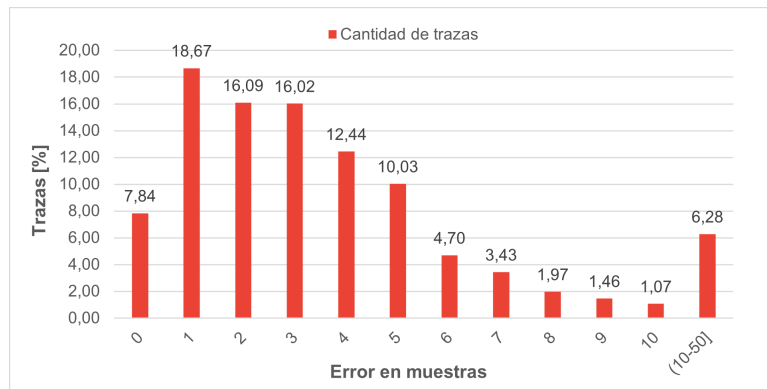
A partir del diseño de red solución se comprueba su desempeño basado en los resultados obtenidos de la predicción del primer arribo. En la Figura 41 se describe el número de trazas cuyo primer arribo fue detectado de forma correcta y el número de trazas con arribos con errores de hasta cinco muestras frente al picado verdadero.

Figura 41.

Histograma del desempeño de la red solución en la predicción de primeros arribos



(a)



(b)

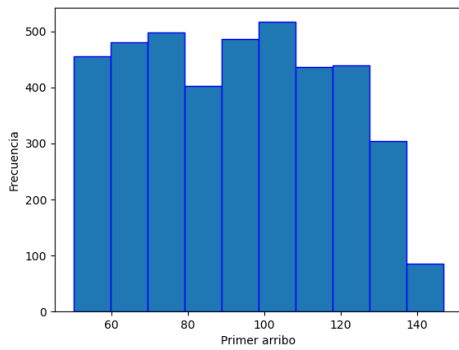
Nota. El conjunto de datos utilizados para validación es de 4108 trazas. De este conjunto de trazas en (a) se ilustra la distribución de arribos para errores en intervalos de 0 a 50 muestras, con un máximo error en la predicción de 46 muestras, y en (b) se muestra a detalle el intervalo de 0 a 10 muestras de la figura a.

La red solución cuenta con valor máximo de FB predicho en la muestra 139 frente a un máximo FB del picado verdadero de 147. El mínimo valor de FB predicho corresponde a la mues-

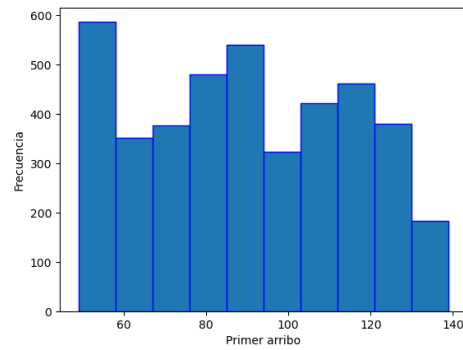
tra 49 frente a la muestra 50 del FB picado verdadero. La distribución de los primeros arribos predichos por la red es comparada frente a la distribución de primeros arribos verdaderos, Figura 42.

Figura 42.

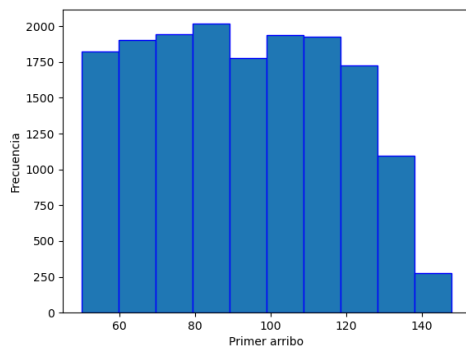
Histograma de distribución de primeros arribos red solución



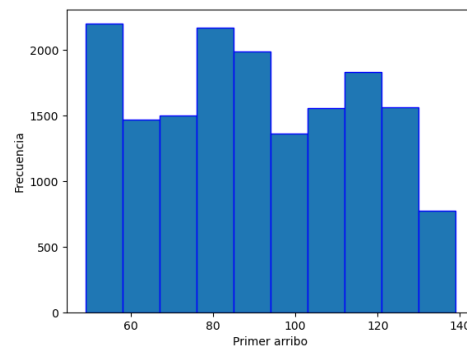
(a)



(b)



(c)



(d)

Nota. (a) Distribución de valores de FB picado verdadero para datos de validación, (b) Histograma de resultados de predicciones de FB para datos de validación, (c) Distribución de valores de FB picado verdadero para datos de entrenamiento y (d) Resultado de predicciones de FB para datos de entrenamiento.

De acuerdo con los resultados obtenidos y partir de consideraciones como lo es que uno de los factores que dificulta la detección de primeros arribos consiste en el ruido que enmascara la traza sísmica y a su vez el FB, se realizó la evaluación de la red solución en presencia de señales con diferentes niveles de ruido. Para esta evaluación, se varió el valor de SNR añadida al conjunto de datos con valores fijos de ruido de 0 dB, 3 dB, 6 dB, 10 dB y 20 dB, para comparar el comportamiento de la red ante trazas con mayores componentes de ruido.

Los resultados de la evaluación del desempeño de la red solución para trazas con mayor presencia de ruido, entrega valores congruentes a pesar del aumento del ruido. En la Tabla 16 determinamos que la red, a pesar de contar con trazas más ruidosas, predice adecuadamente el FB en la mayoría de las trazas.

Tabla 16

Métricas de resultados para red solución con datos de validación con variación de SNR

Relación señal a ruido [dB]	Media	Desviación estándar
0	7	8
3	5	6
6	4	5
10	4	4
20	4	4

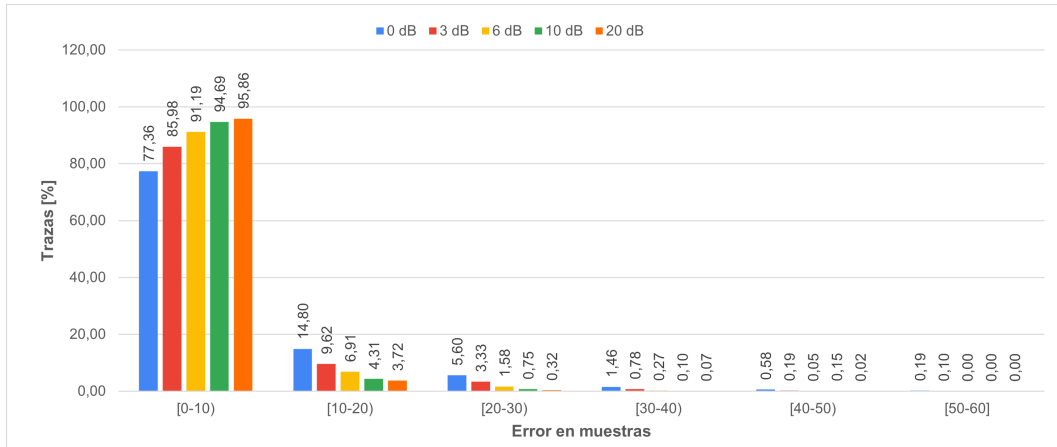
Nota. Las métricas fueron calculadas en términos de muestras para los datos de validación en la red solución. Para el caso de la media, se determinó como el error absoluto entre el resultado del picado verdadero de los datos frente al picado predicho por la red.

De la Figura 43, observamos las diferentes distribuciones de primeros arribos predichos por

la red ante trazas con variaciones de SNR realizadas de forma fija a lo largo de todo el conjunto de datos de validación.

Figura 43.

Histograma de distribución de primeros arribos red solución para datos con SNR variable



(a)



(b)

Nota. El conjunto de datos utilizados para validación es de 4108 trazas las cuales son presentadas a la red bajo SNR variable de 0 dB, 3 dB, 6 dB, 10 dB y 20 dB. En (a) se encuentra la distribución de trazas de acuerdo al intervalo de error en muestras y en (b) la distribución de trazas con errores de una a diez muestras.

La red validada con datos bajo diferentes valores de SNR se comporta con un predicción no superior a las 10 muestras, en donde se encuentra la mayor concentración de trazas. En Figura 43.b, podemos obtener el número de trazas predichas acertadamente y el número de trazas con errores de uno a diez muestras. Para una SNR de 0 dB el máximo error en predicción de la red es de 57 muestras, para una SNR de 3 dB cuenta con un error máximo de 54 muestras, frente a una SNR de 6 dB cuyo error máximo corresponde a 42 muestras, seguidamente una SNR de 10 dB cuenta con un máximo error de predicción de 45 muestras y finalmente un error de 44 muestras para datos con SNR de 20 dB. La red solución cuenta con valor máximo de FB predicho en la muestra 139 frente a un máximo FB del picado verdadero de 147. El mínimo valor de FB predicho corresponde a la muestra 49 frente a la muestra 50 del FB picado verdadero.

Las variaciones de SNR permiten identificar la importancia de considerar el ruido como un factor diferenciador e influyente para la identificación de la red, permitiendo detectar el FB a pesar de tener en mayor o menor proporción ruido en la traza. Con esa idea, decidimos presentarle a la red “trazas” de sólo ruido, las cuales cuentan con una longitud de 1251 muestras. La red solución, identifica y entrega de la predicción de arribos en trazas dentro de los intervalos en los cuales ha sido entrenada y so presentados en la Tabla 17.

Finalmente, dentro de la evaluación del desempeño de la red solución, decidimos presentarle datos de entrada que no se habían utilizado dentro del entrenamiento ni validación de la red (tampoco fueron usados en los experimentos anteriormente presentados). Los datos consisten en las **7223** trazas que fueron descartadas para la implementación de las dos estrategias expuestas en

Tabla 17

Resultados de la predicción de primeros arribos para trazas de sólo ruido.

SNR Aleatorio	FB Predicho
Aleatorio 1	85
Aleatorio 2	81
Aleatorio 3	117
Aleatorio 4	94
Aleatorio 5	70
Aleatorio 6	111

Nota. Los datos son presentados para validar el desempeño de la red solución, cuentan con “trazas” con adición de una SNR aleatoria y la cual comprende es señales de sólo ruido.

el Capítulo 3. Este conjunto de datos cuenta con trazas cuyos valores de FB se encuentran entre la muestra 4 a la 50. A partir de esta característica, las trazas son entregadas a la red solución brindando los resultados presentes en la Tabla 18.

Tabla 18

Métricas de resultados para red solución con datos de validación

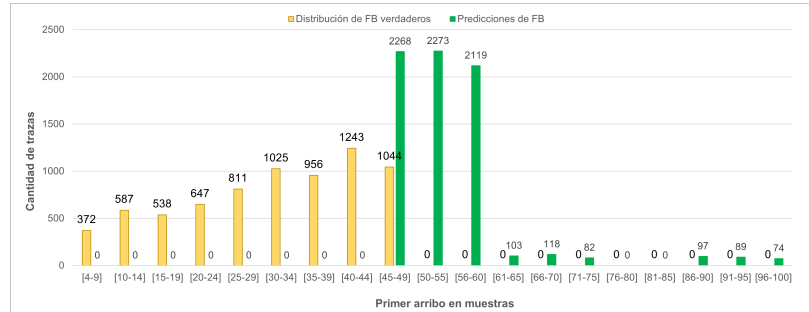
Datos	Media	Desviación estándar
Validación	19	13

Nota. Los datos son presentados para validar el desempeño de la red solución, cuentan con una SNR de 6 a 10 dB como originalmente se planteó la adición de ruido.

En la Figura 44 el máximo FB verdadero en estas trazas es de la muestra 50 y la red predijo en la muestra 139, el mínimo FB verdadero se encuentra ubicado en la muestra 4 y la red lo predice en la muestra 50. Los datos utilizados mantienen la misma SNR de 6 a 10 dB planteada a lo largo del proyecto. La distribución de los primeros arribos predichos por la red son comparados frente a

la distribución de primeros arribos verdaderos.

Figura 44. Histograma de distribución para trazas con primeros arribos de 0 a 50 muestras



Nota. El conjunto de datos de validación corresponden a 7223 trazas con primeros arribos entre las muestras 0 a la 50. Las barras amarillas representan la distribución de los primeros arribos verdaderos sobre el intervalo de 4 a 50 muestras frente a la distribución de primeros arribos predichos por la red representados de color verde.

De la Figura, 44, podemos observar que la red al ser entrenada con valores de FB entre las muestras 50 a la 139, los resultados que entrega se encuentran dentro de ese mismo rango de valores, a pesar de que las trazas que ingresan para la predicción cuentan con valores de FB entre 4 a 50 muestras. La red aprendió de forma adecuada el entrenamiento realizado, sin embargo, este tipo de resultado nos indica que el valor más bajo predicho por la red es el límite inferior aprendido en el entrenamiento de la red, es por ello que la red solución no entrega FB inferiores a la muestra 50.

5. Conclusiones

En este documento se exploraron distintos enfoques en la aplicación de redes neuronales para la detección del primer arribo en un conjunto de trazas sísmicas generadas sintéticamente con características del Valle Medio del Magdalena. Para esto, se diseñó y entrenó una red neuronal con capacidad de detectar el primer arribo ante presencia de niveles de SNR entre 6 a 10 dB, conformada por una secuencia de una capa convolucional, capa de *maxpooling*, capa densa, capa *flatten* y tres capas densas (Figura 36), entregando aciertos en un 7,84 % del conjunto de 4108 trazas de validación, frente a los resultados obtenidos con la implementación de las redes neuronales profunda y convolucional, con 9,07 % y el 3,33 % de aciertos respectivamente. Si bien, los aciertos entregados por la red densamente conectadas son superiores frente a la red solución, cabe recalcar que la red densamente conectada cuenta con predicciones superiores al rango de valores verdaderos. Para el caso de la convolucional, la predicción con errores nulos es baja, por lo cual decidimos unir las dos arquitecturas para mejorar la predicción de primeros arribos, lo que se logra constatar por la media de error en cada red, DNN con media de 65 muestras, CNN con media de 54 muestras y red solución con media de 4 muestras.

La adición de ruido (SNR) y la normalización a las trazas sísmicas dentro del conjunto de datos, empleados para la implementación de las estrategias tomadas del estado del arte y del diseño

de la red solución, son un factor fundamental en el entrenamiento, debido a que la adición de ruido permite simular un comportamiento sísmico más real y la normalización proporciona uniformidad en las amplitudes de la traza. Su importancia es determinada a partir de resultados obtenidos en la primera estrategia, debido a que los datos sin ruido facilitan la identificación del primer arribo, al mecanizar de forma casi trivial que durante el entrenamiento, el cambio de amplitud consistía en el primer arribo, por lo cual durante los primeros tramos de la traza sísmica las variaciones no eran perceptibles ni considerables en la propagación de la onda refractada y la primera variación notoria de amplitud correspondía al primer arribo (pasar a un valor diferente de cero en la magnitud del campo).

Para el diseño de la red solución, se destaca la importancia de la selección de parámetros e hiperparámetros, los cuales contribuyen a mejorar el desempeño de la red de acuerdo, al ajuste de los mismos frente al conjunto de datos con los que se contaba durante las distintas implementaciones. De la misma forma, durante el entrenamiento la presentación de datos en la red neuronal juega un papel consecuente en la implementación debido a que, a pesar de que las dos redes cuenten con exactamente la misma arquitectura, como es el caso de las redes DNN pre y pos FB, los resultados son diferentes debido a que los datos de entrada corresponden a ventanas temporales con diferentes enfoques de la misma traza, lo que implica un entrenamiento bajo características diferentes.

Para el diseño de la red solución, se destaca la importancia de la selección de parámetros e hiper parámetros, los cuales contribuyen en el mejoramiento del desempeño de la red, debido a la mejor adaptación dentro del modelo entrenado, esto se visualiza por medio de las métricas

obtenidas en cada implementación, al ser notorio el cambio del error.

En la primer estrategia las redes DNN durante el entrenamiento son sensibles a la forma y orden en que se presentan segmentos de trazas sintéticas. A pesar de que las dos redes tienen exactamente la misma arquitectura (red pre y red pos FB), los resultados son diferentes debido a que los datos de entrada corresponden a ventanas temporales con distinta información de la misma traza, lo que implica un entrenamiento bajo características diferentes y provoca dos redes distintas.

Los resultados obtenidos a partir de la función de pérdida y validación deben estar alineados con la salida presentada por la red, si la función de pérdida es baja nos indica que la salida de la red realiza una buena aproximación frente a las etiquetas utilizadas para el entrenamiento. Sin embargo, para salidas cuyo comportamiento no es ideal, la salida de la red entrega información suficiente para inferir, mediante el posprocesamiento de la primera diferencia la identificación del primer arribo, y se esperaría que para una menor función de pérdida el posprocesamiento debe igualar, por lo menos, a la predicción que tenemos sin posprocesamiento. La función de pérdida y validación durante el entrenamiento busca aumentar el aprendizaje de la red, mejorando la generalización en la salida. Esto se logra visualizar en las implementaciones presentadas en la sección 4.1, las cuales arrojaron salidas diferentes a la etiqueta utilizada en el entrenamiento. Con estos resultados, las curvas de la función de pérdida entregaban valores no congruentes y fue necesario establecer que los criterios de selección se realizaban a partir de la salida de la red y las métricas de predicción.

La aplicación de transformaciones (e.g. entropía, wavelet) permite tener información más diversa de la traza bajo diferentes dominios que son entregados a la red para contribuir en el entrenamiento de la misma, al contar con más datos de los cuales la red pueda obtener e identificar características y patrones en la detección de primeros arribos. Dentro del uso de transformadas para el entrenamiento de la red, para el caso del uso de la transformada Wavelet, empleada en la implementación convolucional de la segunda estrategia, el primer arribo tiene unas características que pueden ser reflejadas en la descomposición wavelet, debido a que podemos proceder a un análisis del comportamiento de la traza bajo los criterios de frecuencia y escala, permitiendo así extraer un mayor número de características. Adicionalmente, el análisis de wavelets se usa habitualmente para señales no periódicas, como las trazas sísmicas. Dentro de las ventajas para el uso de transformadas en el dominio de la frecuencia, el aporte en frecuencia por parte de Fourier entrega muy poca información, al perder casi toda la información temporal de la traza, pero su utilización permite determinar las frecuencias máximas de corte en el que se concentra el primer arribo. Frente a esto, las wavelets permiten detectar singularidades de forma más eficiente a pequeñas escalas.

El rango de primeros arribos verdaderos utilizados durante el entrenamiento condiciona las predicciones realizadas a la salida de la red durante la detección de primeros arribos en datos de validación. Esto se comprueba con los resultados presentados en la Figura 44, en donde a la red son presentadas trazas con características particulares a las empleadas en el entrenamiento. Este conjunto de datos indica que la red no responde de forma adecuada frente a trazas cuyos primeros arribos no se encuentren en el rango empleado en el entrenamiento, debido a que entrega falsas

predicciones en muestras entre la 49 a la 100, frente al intervalo verdadero de la muestra 4 a la 49.

Para trazas con bajas componentes de ruido, la red solución responde con un alto porcentaje de acierto en la detección de primeros arribos. Esto se logra comprobar, a partir de los resultados presentados en la Figura 43, nos permiten observar la distribución de trazas frente a variaciones de SNR para el conjunto de validación; logramos observar que entre menos ruido contenga la traza disminuye el error en predicción al tenerse una mayor concentración de trazas en intervalos con errores de 0 a 9 muestras.

El diseño de una red neuronal no tiene fronteras, se pueden abordar diferentes ideas y perspectivas hasta donde el contexto, tiempo y recursos disponibles para la solución del problema lo permita. Por lo cual, se concluye que el diseño de nuestra red neuronal no tiene una solución final, por el sin número de combinaciones posibles, lo cual hace que siempre pueda existir una mejor red ante el problema de primeros arribos. Teniendo presente que en gran parte la red (organización de datos, procesamiento y criterios de selección en parámetros e hiperparámetros) es el resultado de nuestra perspectiva como diseñadores.

6. Recomendaciones

- Con propósito de generalizar la red solución expuesta en el Capítulo 4, para realizar predicciones a trazas con distintas cantidades de muestras, se plantea hacer el entrenamiento de la red, donde el etiquetado corresponda a una máscara binaria, la cual cuenta con valores de

uno en una ventana con longitud previamente definida cuyo centro es el primer arribo, la ventana cuenta con valores de uno y el resto de las muestras de la traza fuera de la ventana con valores de cero. La salida de la red deberá entregar de forma ideal, la ventana con valores uno y el valor central corresponderá al FB predicho.

- La aplicación de diversas transformaciones a la traza sísmica permite adquirir nuevas características del FB que son visibles bajo diferentes dominios. La adición de nuevas transformaciones (por ejemplo, la entropía) a la red solución puede contribuir a la mejora en el proceso de predicción, siempre y cuando, se mantenga la forma y el orden en el que son presentados los datos a la red.

- El uso de otro tipo de ondícula madre para la transformada wavelet, permitiría mejorar los tiempos de predicción y la precisión de la detección, por lo cual se invita a realizar variaciones sobre la ondícula madre, una wavelet comúnmente utilizada en ingeniería sísmica es la conocida como pulso de *Ricker* o sombrero mexicano, esta wavelet no tiene ningún escalamiento y es derivada desde una función que es proporcional a la segunda derivada de una función Gaussiana.

Referencias Bibliográficas

- Allen, R. (1978). Automatic earthquake recognition and timing from single traces. *Geophysics*, 68(5):1521–1532.
- Barnes, A. E. (2007). A tutorial on complex seismic trace analysis. *Geophysics*, 72(6):W33–W43.
- Barnes, A. E. (2016). Handbook of poststack seismic attributes. *Society of Exploration Geophysicists*.
- Barráez, C. E. G. (2000). Detector de eventos sísmicos en tiempo real utilizando redes neuronales. *Interciencia*, 25:293–298.
- Bogiatzis, P. e Ishii, M. (2015). Continuous wavelet decomposition algorithms for automatic detection of compressional- and shear-wave arrival times. *Bulletin of the Seismological Society of America*, 105(3).
- Boschetti, F., List, R., & Dentith, M. (1996). A fractal-based algorithm for detecting first arrivals on seismic traces. *Geophysics*, 61:1095–1102.
- Brown, A. R. (2011). Interpretation of three-dimensional seismic data. *Society of Exploration Geophysicists and American Association of Petroleum Geologists*.

- Chopra, S. & Marfurt, K. J. (2007). Seismic attributes for prospect identification and reservoir characterization. *Society of Exploration Geophysicists and European Association of Geoscientists and Engineers*.
- Clement, A. & Peraldi, R. (1972). Digital processing of refraction data study of first arrivals. *Geophysical Prospecting*, 20:529–548.
- CodeTD (2020). Biblioteca de aprendizaje profundo basada en keras python. url: <https://www.codetd.com/es/article/10901016>.
- Coppens, F. (1985). First arrival picking on common-offset trace collections for automatic estimation of static corrections. *Geophysics*, 33:1212–1213.
- Córdova, N. C. (s.f.). El principio de entropía. pp. 2–3. https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios_catedras/obligatorias/056_adolescencia2/material/fichas/el_principio_de_entropia.pdf.
- Denis, A. & Crémoux, F. (2002). Using the entropy of curves to segment a time or spatial series. *Mathematical Geology*, p. 34. 899–914.
- Earle, P. & Shearer, P. (1994). Characterization of global seismograms using an automatic-picking algorithm. *Geophysics, Bulletin of the Seismological Society of America*, 2(84):366–376.
- Gelchinsky, B. & Shtivelman, V. (1983). Automatic picking of first arrivals and parameterization of traveltimes curves. *Geophysical Prospecting*, (31):915–928.

Gillfeather, T., Horrocks, T., Holden, J., & Wedge, D. (2021). A comparative study of neural network methods for first break detection using seismic refraction data over a detrital iron ore deposit. *Geophysics*, 137(01):2–6.

Google (s.f.). Colaboratory. <https://research.google.com/colaboratory/>.

Ibarra, O. E. C. (Bogotá, Colombia, 2011). Detección y clasificación automática de registros sísmicos en el observatorio vulcanológico y sismológico de pasto utilizando redes neuronales artificiales (tesis de maestría.). *Universidad Nacional de Colombia Facultad de Geociencias, posgrado Geofísica*.

Jupyter (2022). Jupyter notebook. url: <https://jupyter.org/>.

Kaus, B. (2013). Project: Elastic wave propagation in 2-d using a staggered grid method. *Gutenberg University*.

Krohn, C. E. (1984). Geophone ground coupling. *Geophysics*, 49(6):722–731.

Lara, A. (2015). Descripción del proceso analógico y conversión a digital, sísmico en la exploración de hidrocarburos. *Tesis (Ingeniería en Comunicaciones y Electrónica), Instituto Politécnico Nacional, ESIME, Unidad Culhuacan*, p. 107 páginas.

Larranaga, P., Inza, I., & Moujahid, A. (2007). Redes neuronales. *Departamento de Ciencias de la Computación e Inteligencia Artificial*, p. 19.

McClelland, J. & Rumelhart, E. (1986). Parallel distributed processing. *Geophysics*, Vol. 2.

Mexicano, S. G. (s.f.). *¿Qué es la Geofísica?* pp. 1–2. <https://www.sgm.gob.mx>.

Mustafa Senkaya, H. K. (2014). A semi-automatic approach to identify first arrival time: the cross-correlation technique (cct). *Earth Sciences Research Journal (ESRJ)*, 18(2):107–113.

Męzyk, M. & Malinowski, M. (2019). Multi-pattern algorithm for first-break picking employing open-source machine learning libraries. *Journal of Applied Geophysics*, 170:103848. <https://www.sciencedirect.com/science/article/pii/S0926985119302435>.

Ruiz, R. O. (México, 1994). Reconocimiento de formas de onda en registros sísmicos mediante redes neuronales (tesis de maestría.). *Instituto Tecnológico y de Estudios Superiores de Monterrey*.

Rumelhart, D., Hinton, G., & Williams, R. (1986). Learning representations by backpropagation errors. *Geophysics*, pp. 533–536.

Sabbione, J. I. (2012). Algoritmos matemáticos y computacionales para la detección automática de señales sísmicas. *Facultad de Ciencias Astronómicas y Geofísicas*, pp. 19–21.

Sabbione, J. I. & Velis, D. (2010). Automatic first-breaks picking: New strategies and algorithms. *Geophysics*, 75(4):V67–V76.

Saragiotis, C., Hadjileontiadis, L., Rekanos, I., & Panas, S. (2004). Automatic p-phase picking using maximum kurtosis and k-statistics criteria. *IEEE*, 1(3):147–151.

Segyio (2022). Python segyio library.

Shtivelman, V. & Gelchinsky, B. (1983). Automatic picking of first arrivals and parameterization of traveltime curves. *Geophysical Prospecting*, 31:915–928.

Servicio Geológico Colombiano (2002). Manual de entrega de información técnica y geológica de las actividades de evaluación, exploración y producción de hidrocarburos al banco de información petrolera – anexo 1. información geofísica. https://www2.sgc.gov.co/Noticias/tpaginas/Entrega-Informacion-BIP/Manual_V0_AnexoNo-01_Informacion-Geofisica_V1.0_Mar19_2020.pdf.

Spagnolini, U. (1991). Adaptive picking of refracted first arrivals. *Geophysical Prospecting*, 31:293–313.

Sánchez, A. O. (México, 2006). Algoritmo para detección de vibraciones anormales en maquinarias utilizando la transformada wavelet (tesis de maestría.). *Departamento de Computación, Electrónica, Física e Innovación. Escuela de Ingeniería y Ciencias, Universidad de las Américas Puebla.*, p. Capítulo 2.

Tensorflow (2021). *Plataforma de extremo a extremo de código abierto para el aprendizaje automático.*, definición. Disponible en: .

Tibuleac, I. M., Herrin, E. T., Britton, J. M., Shumway, R., & Rosca, A. (2003). Automatic secondary seismic phase picking using wavelet transform. *Proceeding of 25th Seismic Research Review (SRR-03)*, pp. 352–359.

Toro, L. (2022). *Anaconda, ciencia de datos con Python*, definición. Disponible en:
<https://blog.desdelinux.net/ciencia-de-datos-con-python/>.

Torres, J. (Barcelona, 2018). Deep learning, introducción práctica con keras, primera parte. *Independently published*.

Tsai, K. C., Hu, W., Wu, X., Chen, J., & Han, Z. (2020). Automatic first arrival picking via deep learning with human interactive learning. *IEEE Transactions on Geoscience and Remote Sensing*, 58(2):1380–1391.

Wavelet (2022). Python wavelet transform library. *Signals*.

Yung, S.K., L. T. I. (1997). An example of seismic time picking by third-order bicoherence. *Geophysics*, 6(62):1947–1951.

Apéndices

A. Códigos de programación implementados en estrategias y red solución

El código referente a cada uno de los prototipos, solución, y resultados expuestos a lo largo de todo el documento, fueron desarrollados por los autores, junto con la orientación del director y codirector del presente proyecto.

El desarrollo del proyecto, se encuentra disponible para su revisión en el repositorio de GitLab del Grupo de Investigación en Conectividad y Procesamiento de Señales (CPS), en donde se encuentra el resultado de implementar todas las pruebas y el diseño de la red catalogado como solución. Dicho repositorio se encuentra organizado por cinco diferentes carpetas:

- Carpeta ubicada en fb-seismic/DNN2FINAL contiene la implementación de la red neuronal profunda densamente conectada expuesta por el autor (Mężyk y Malinowski, 2019).
- Carpeta ubicada en fb-seismic/snr610_mse_cnn33_500_original contiene la implementación de la red convolucional expuesta por los autores (Gillfeather *et al.*, 2021).
- Carpeta ubicada en fb-seismic/RES_PruebasCNN contiene la implementación de las pruebas descritas en la Tabla 7, realizadas con base en las redes expuestas por Mężyk y Malinowski

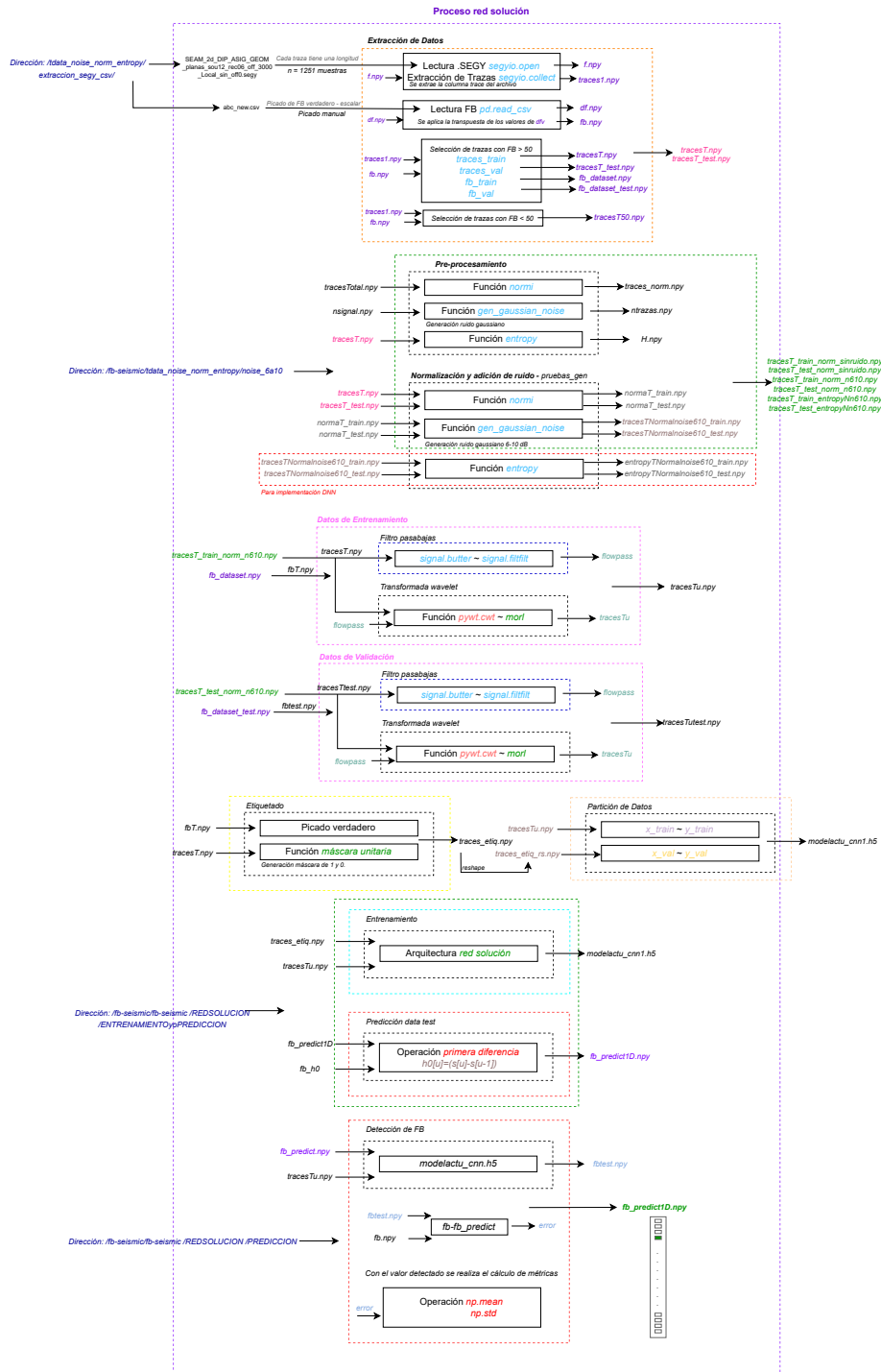
(2019); Gillfeather *et al.* (2021).

- Carpeta ubicada en `/fb-seismic/tdata_noise_norm_entropy` se encuentran el dataset usado en todas las implementaciones.
- En la carpeta ubicada en `fb-seismic/config_dense_cnnn_BN_FINAL` contiene el diseño expuesto como solución respecto a las trazas con características del valle medio del Magdalena usadas por los autores del presente proyecto.

Durante todo el proyecto se maneja un mismo entorno de Anaconda para realizar y ejecutar todas las implementaciones, en particular se trabajó con la versión de Anaconda 2021.05 con Python 3.8.8. Para obtener una réplica de los archivos que contienen el desarrollo de todo el proyecto y por consiguiente obtener los resultados aquí expuestos, se recomienda trabajar con el gestor de entornos Anaconda y hacer uso de la herramienta git, ejecutando el comando `git clone` seguido del link del repositorio https://gitlab.com/CPS_Repo/Seismic/semillero2020/fb-seismic.git

En la Figura 45 se realiza un diagrama de flujo en el que se observa la cadena de procesamiento y las funciones empleadas en cada etapa, se indica lo que se realiza desde que se recibe un archivo SEG-Y, hasta que es presentado el dato detectado como primer arribo.

Figura 45.
Diagrama de flujo del procesamiento de la detección de primeros arribos



Nota. Se describe el procesamiento realizado dentro del código implementado para la detección de primeros arribos. Para el entrenamiento, se requiere que los datos se encuentren organizados en matrices, de tal forma que cada archivo `.npy` (archivo que contiene las trazas y otro con el picado) la información debe estar ubicada en una columna por traza. Para la predicción, los datos de las trazas están organizado en columnas.