

# A Deep Learning Approach for Epilepsy Seizure Identification using Electroencephalogram Signals: A Preliminary Study

Sergio Jacobo-Zavaleta , Jorge Zavaleta 

**Abstract**—Epilepsy is a neurological disease that affects around 50 million people of all ages worldwide. In this study, five deep learning networks were compared to determine the best performance in seizure detection using electroencephalogram raw signals from the TUH EEG Seizure Corpus database. The methodology included three strategies for reducing the high computational cost of training time series data: Extracting epileptic features from patient signals by concatenating all seizure events into a shorter single one, selecting signals of duration greater than 180 seconds, and generating two randomized groups based on patient and non-patient (control) signals for larger and shorter supervised training-validation processes. Finally, the models were evaluated using two groups, one is formed of patient-control data and the other using only patient data. The results showed that a simple LSTM-based network, a hybrid network and the reported ChronoNet achieved the best metric performance for a binary classification, with up to 71.50 % of sensibility and 83.70 % of specificity for a patient-control detection; and up to 56.60 % of sensibility and 95.90 % of specificity for a patient-specific detection. In conclusion, deep learning-based models might automate seizure detection in order to improve epilepsy diagnosis and accelerate early treatments using electroencephalogram signals.

**Index Terms**—Deep learning, Electroencephalogram, Epilepsy, Seizure

## I. INTRODUCCIÓN

La epilepsia es una enfermedad neurológica crónica que afecta cerca de 50 millones de personas en el mundo de todas las edades [1]. La actividad cerebral anormal que se produce conlleva a convulsiones involuntarias y recurrentes en distintas regiones cerebrales de forma focalizada o generalizada (conocida como crisis), dañando muchas veces la conciencia de los pacientes [2]. Esto conlleva a diversas limitaciones en su vida diaria, incluido el frecuente estigmatismo social [3]. Reduciendo así, la calidad de vida del 80 % de pacientes debido a la depresión, ansiedad y aislamiento social [4].

En adultos, la etiología epiléptica está asociada a múltiples factores de riesgo como la predisposición genética, metabólica, infecciosa, inmunidad y tumores cerebrales [5], agravado por la aparición de desórdenes degenerativos durante la vejez [6]. Además, existe un riesgo de muerte prematura de hasta tres veces más que la población general, debido al estado epiléptico repentino de las crisis que derivan en mayor trauma cerebral y

accidentes domésticos [7], pudiendo desarrollarse en cualquier momento de sus vidas [8]. No obstante, el 50 % de los casos son idiopáticos o permanecen desconocidos [9].

En Latinoamérica, la epilepsia es problema de salud pública que trae como consecuencias el registro de pocos datos epidemiológicos y pacientes tardíamente diagnosticados. La prevalencia es variable, alcanzando valores de hasta 139 por cada 100 000 personas en países de medianos y bajos ingresos, frente a un 48.9 en países de altos ingresos. Agravado posiblemente por el riesgo de enfermedades endémicas como el paludismo y neurocisticercosis en la región [4]. No obstante, el uso correcto de medicinas y tratamientos oportunos podrían reducir hasta un 70 % los casos hospitalarios [10].

En la práctica clínica, la electroencefalografía es la principal herramienta diagnóstica de epilepsia al medir la actividad eléctrica a nivel cerebral utilizando electrodos ubicados en el cuero cabelludo o directamente en la superficie del cerebro. Mediante el monitoreo y grabación de señales de electroencefalograma (EEG), los neurólogos pueden analizar e identificar actividad epileptiforme, ocurrencia de convulsiones y localizar zonas afectadas, para brindar mejores tratamientos y decisiones terapéuticas antes de una cirugía [11]. Sin embargo, el análisis visual y tareas manuales relacionadas con la identificación de patrones epilépticos son extremadamente lentos (de horas a días), laboriosos, costosos y asociado a una demanda de neurólogos altamente entrenados [12]. Desventajas que han motivado la búsqueda de soluciones tecnológicas para la detección automática de convulsiones y mejorar la calidad de vida de pacientes epilépticos mediante el análisis de señales EEG, debido a su alta resolución temporal, evaluación no invasiva y relativo bajo costo [13]. Impulsado por nuevas formas de grabación portable e inalámbrica de la señal [14].

Por otro lado, el desarrollo y maduración de múltiples tecnologías digitales han permitido afrontar grandes problemas clínicos y enfermedades en los últimos años [15], como el monitoreo de riesgos asociados a la actividad física en pacientes obesos [16] o el diagnóstico de enfermedades neurodegenerativas a partir del análisis de imágenes [17]. Entre ellas, destaca el desarrollo de técnicas en *Machine Learning* (ML) y *Deep Learning* (DL) como subcampos de la inteligencia artificial para aplicaciones de neurociencia e ingeniería neural. Este último, basado en modelos de redes neuronales artificiales para resolver problemas de clasificación de eventos como convulsiones epilépticas. No obstante, para ello es necesario afrontar dos retos ampliamente documentados. Primero, la significativa variabilidad de señales EEG en diferentes indi-

Sergio Jacobo-Zavaleta is with Universidad Nacional de Trujillo, e-mail: sjacobo@unitru.edu.pe

Jorge Zavaleta is a Postdoctoral Research at at the State University of Rio de Janeiro (UERJ) in the project CAPES-Telemedecine and Medical Data Analysis, e-mail:zavaleta@pet-si.ufrj.br

viduos e incluso, en el mismo paciente. Y segundo, la alta susceptibilidad de las señales frente la presencia de artefactos fisiológicos causados por movimientos involuntarios y ruido no fisiológico. Existiendo muchas veces una semejanza entre señales convulsivas y normales [18].

Entre las principales redes neuronales, destacan las redes convolucionales (CNN), recurrentes (RNN), recurrentes con *Long Short-Term Memory* (LSTM) y recurrentes con *Gated Recurrent Units* (GRU) [19]. Que junto a la creciente disponibilidad de grandes bases de datos públicas, han acelerado la implementación de diversos sistemas computacionales para la detección, diagnóstico y pronóstico automático de epilepsia antes de una crisis convulsiva [20], [21].

Por lo anterior, este trabajo evaluó el rendimiento de múltiples arquitecturas de redes neuronales para la identificación de epilepsia mediante la clasificación binaria de señales convulsivas y no convulsivas, extraídas de la base de datos más diversa públicamente disponible para señales crudas EEG. De poder generalizarse, estos modelos podrían formar parte de un mayor sistema con gran potencial de herramienta diagnóstica y pronóstica a corto plazo, al acelerar la interpretación de largos volúmenes de datos, disminuir costos y aligerar la carga de trabajo frente a una baja disponibilidad de expertos certificados en el diagnóstico de enfermedades neurológicas.

El resto del artículo está organizado como sigue. La sección 2 describe los estudios relevantes más recientes para la clasificación automática de convulsiones utilizando DL, particularmente enfocadas en redes LSTM e híbridas. La sección 3 describe la metodología para el modelamiento, entrenamiento, validación y evaluación de las redes neuronales. Los resultados se discuten en la sección 4. Finalmente, las conclusiones y observaciones futuras se describen en la sección 5.

## II. TRABAJOS RELACIONADOS

La detección de convulsiones y no convulsiones es un problema de clasificación de señales a partir de segmentos de datos EEG. Para ello, numerosos estudios se han realizado en las últimas dos décadas con resultados prometedores. Desde técnicas clásicas y novedosas en ML, hasta las más recientes basadas en DL como el uso de redes neuronales artificiales. Así, las redes neuronales convolucionales CNN se han utilizado para el análisis bidimensional (2D) de señales EEG representadas en imágenes para capturar patrones espaciales y espectrales [12], [18], [22] con buenos resultados. No obstante, la naturaleza dinámica de tipo unidimensional (1D) de las señales han llevado su abordaje temporal con datos de longitud variable mediante el uso de redes 1D-CNN [23], RNN [24], LSTM, GRU e híbridos [19]. A continuación, se describen estos estudios publicados en los últimos cinco años.

Golmohammadi et al. [25] evaluaron dos arquitecturas LSTM con tres y cuatro capas con la función de activación Softmax para una clasificación multiclase de señales EEG utilizando por primera vez los datos del TUH EEG Seizure Corpus, con sensibilidades en el rango de 30%. Más tarde, Talathi utilizó una red con cuatro capas de tipo GRU para la detección de convulsiones, alcanzando una exactitud de 98% dentro de los primeros cinco segundos de una crisis con datos

provistos por la University of Bonn. Considerada como una de las bases de datos disponible más antigua [26]. De igual modo, Chen et al. utilizaron los mismos datos para entrenar varias redes basadas en 1D-CNN, LSTM y GRU, alcanzando exactitudes de 97.27%, 96.82% y 96.67% respectivamente [27]. Por su parte, Hussein et al. diseñó una red neuronal profunda de tipo LSTM con una cantidad de entradas igual al número de canales EEG, reportando un 100% de sensibilidad, especificidad y exactitud [28].

Más tarde, Fukumori et al. implementaron una red automática de bandas de frecuencia utilizando capas CNN-RNN (hasta 31) para la detección de picos epilépticos en señales EEG [29]. Alcanzando métricas de hasta 88.80%, 88.60% y 88.69% de sensibilidad, especificidad y precisión respectivamente, mediante el uso de una nueva arquitectura para superar la variabilidad morfológica de las convulsiones y problemas asociados a la profundidad de redes RNN [30].

Por otro lado, el número de capas de red no garantiza un mejor rendimiento, como el uso de 7 capas (99.16% de exactitud) [31] o 14 capas (90.60% de exactitud) [32]. Sin embargo, existen claras diferencias debido al tipo de preprocesamiento utilizado. Mediante la *Stockwell's transformation* como método de extracción de características, una red LSTM bidireccional alcanzó un 98.09% de sensibilidad y 98.68% de especificidad [33]; utilizando la *discrete wavelet transform*, otra red alcanzó un área bajo la curva de Característica Operativa del Receptor (ROC) superior a 0.79 [29]; o aplicando *Wavelet Filtering*, se alcanzó un 98.50% de exactitud [34].

En la mayoría de los casos, se utilizaron sesiones EEG de un solo canal, proveniente de bases de datos ampliamente conocidos, muchos de ellos sin presencia de artefactos ni ruido. Lo que lleva a una baja generalización o robustez de los modelos. A ello se suma, la extensa diversidad de técnicas manuales utilizadas para la extracción y selección de características, llevando a una dependencia de los datos de entrada y selección de hiperparámetros durante el preprocesamiento [23].

## III. PROPUESTA DE SOLUCIÓN

Dado que la epilepsia es un desorden heterogéneo frecuentemente asociado con hallazgos anormales de señales EEG en la práctica clínica, muchos métodos han sido propuestos para clasificarlas de forma binaria (señal normal vs. epiléptica o señal convulsiva vs. no convulsiva), terciaria o multiclase. No obstante, para demostrar el impacto del uso de señales crudas en el rendimiento de modelos de DL, se evaluaron diversas arquitecturas de redes neuronales simples e híbridas para la clasificación binaria de segmentos convulsivos.

### A. Selección de los Datos

Para representar la gran variabilidad de señales epileptiformes entre individuos, se utilizaron datos de la Temple University Hospital (TUH) EEG Seizure Corpus (TUSZ) (versión 1.5.4) [35], un subconjunto del TUH Abnormal EEG Corpus [36]. TUSZ contiene el historial de grabaciones EEG para 675 personas con un total de 1 643 sesiones (1 475 horas), de las cuales 528 sesiones contienen convulsiones. Entre sus principales ventajas sobre otras bases de datos, destacan su

extensa diversidad morfológica de convulsiones clínicamente etiquetadas, duraciones, frecuencias de muestreo, presencia de artefactos, amplia distribución etaria y datos agrupados en conjuntos de entrenamiento, validación y evaluación en archivos *European Data Format* (EDF). Ideal para el estudio de modelos supervisados con DL en problemas de clasificación binaria o multiclase [36]. Donde, modelos entrenados con eventos morfológicamente diversos tienen el potencial de reducir significativamente el número de falsos positivos [25].

**B. Preparación de los Datos**

**B1. Formulación de los Datos de Entrada:** Para no limitar las sesiones a una longitud máxima y aprovechar la mayoría de señales convulsivas, se realizaron dos estrategias para reducir sus largas duraciones (muchas de ellas grabadas por horas con una frecuencia de muestreo de 250 Hz y 19 electrodos) según la Fig. 1. Primero, utilizando las anotaciones diagnósticas para cada archivo, se identificaron los periodos convulsivos para luego realizar múltiples segmentaciones de la señal original de acuerdo a las etiquetas temporales por cada sesión; y segundo, todas las partes segmentadas se concatenaron en orden y sin superposición para formar una nueva señal de menor duración y almacenamiento.

Luego, mediante la visualización de un histograma para la distribución de duraciones (sesgado a la derecha) de todas las nuevas señales, se seleccionaron todas aquellas con una duración menor al de la mediana (180 segundos) a fin de uniformizar la longitud de los datos de entrada por mayoría.

Por otro lado, para el entrenamiento de las redes basadas en LSTM fue necesario convertir las señales 2D anteriores (con tamaño matricial de pasos de tiempo x número de canales) en datos 3D, mediante su partición en una secuencia de muestras o segmentos de igual duración y sin superposición (número de muestras x pasos de tiempo para la duración fija x número de canales). La Fig. 2 muestra esta representación obtenida con ayuda de la librería de Python MNE (versión 1.2) para la exploración y análisis de bioseñales.

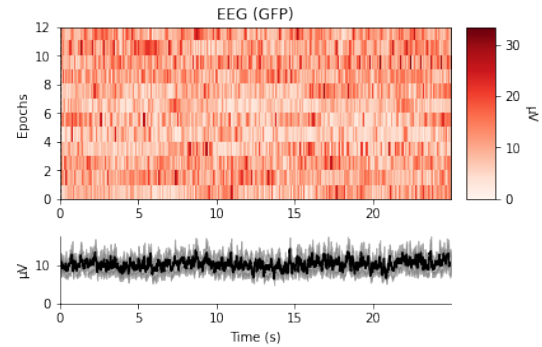


Fig. 2. Visualización 3D de señales EEG para un paciente epiléptico. Donde *epochs* es el número de muestras.

**B2. Datos de entrenamiento, validación y evaluación:** Para minimizar el costo computacional de los entrenamientos, se prepararon dos conjuntos de datos utilizando el algoritmo de la Fig. 3. Mediante una división aleatoria 70-30 del total de registros previamente obtenidos, del grupo grande (color amarillo) se obtuvieron 32 sesiones para entrenamiento y 14 para validación. Sin embargo, las 14 sesiones de entrenamiento del grupo reducido (color azul) se validaron también con las 14 sesiones anteriores, ignorando las 6 sesiones correspondientes (30% de 20) por considerarse una cantidad muy pequeña.

Por su evaluación, se escogieron dos grupos con igual cantidad de sesiones por caso de estudio según la Fig 4. Los datos con ausencia de convulsiones se definieron como casos de control por su rol de comparación. El primero, basado en la combinación de datos previamente descritos y el segundo, utilizando únicamente datos de pacientes mediante la extracción y concatenación de sus periodos convulsivos y no convulsivos (emulando casos de control). Este último, con la finalidad de realizar una detección específica de pacientes epilépticos [12]. Finalmente, los datos de entrenamiento (con estructuras matriciales 3D) fueron balanceadas de acuerdo con el número de etiquetas (0:no convulsión, 1:convulsión) utilizando una técnica de *oversampling*.

**B3. Preprocesamiento de la señal:** Físicamente, las sesiones fueron grabadas bajo la configuración estándar de electrodos 10-20, utilizando al menos 19 electrodos para 22 canales EEG en un montaje bipolar. Ideal para minimizar el ruido durante la grabación [35].

Para la eliminación de ruido y artefactos en cada canal de la señal, se seleccionó un filtro pasabanda con frecuencias de 1-70 Hz, utilizado en muchos estudios [13]. Seguido por una normalización estándar de los datos antes del entrenamiento.

**C. Arquitectura de las Redes**

En los últimos años, el éxito de las redes CNN para resolver problemas de clasificación de convulsiones epilépticas ha sido generalmente llevado al estudio de señales EEG como datos de entrada bidimensional en el dominio de la frecuencia. No obstante, el análisis de datos y señales como series de tiempo han llevado al desarrollo de redes RNN, 1D-CNN e híbridas con diferente número de entradas y salidas [13].

De esta manera, para evaluar el impacto de las redes sobre los datos, se implementaron tres redes simples con un

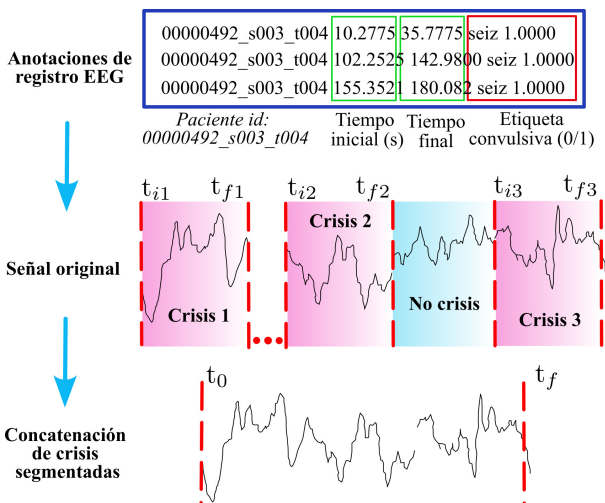


Fig. 1. Segmentación temporal de una señal EEG en periodos convulsivos utilizando etiquetas para datos del TUSZ.

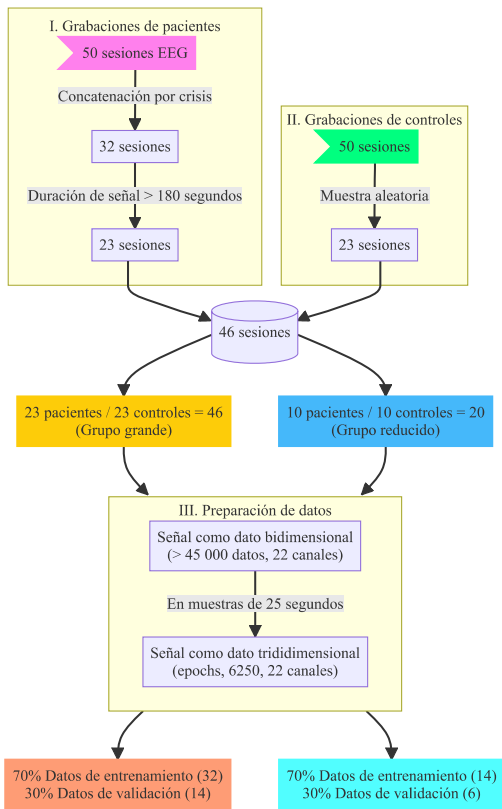


Fig. 3. Preparación de datos de entrenamiento y validación.

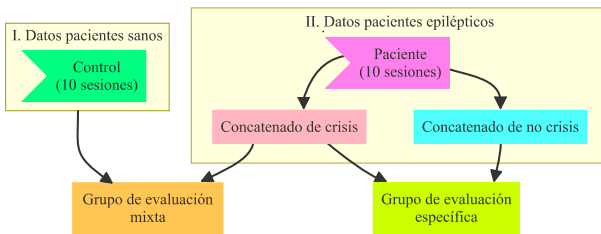


Fig. 4. Preparación de datos de evaluación.

incremento de capas entre sí (1, 2 y 4) y dos más complejas previamente reportadas (3 y 5) según la Tabla I, con la finalidad de evaluar el rendimiento de estos modelos para los datos TUSZ, utilizando principalmente la librería TensorFlow (versión 2.9.2) para la implementación de redes neuronales.

#### IV. RESULTADOS

Para evaluar la efectividad de los diferentes modelos de redes neuronales en la detección de convulsiones, comparamos sus rendimientos con los obtenidos por otros modelos basados también en LSTM e híbridos que utilizaron la misma base de datos o derivados de ella. Este rendimiento de detección fue evaluado usando las métricas estándares de sensibilidad (*recall*), especificidad, exactitud, precisión, y área bajo la curva ROC (AUC). Resumidas en las Tablas II, III y IV.

Para el entrenamiento de las redes se utilizaron los siguientes parámetros base: *batch size* de 32; una función pérdida de error (diferencia entre los valores predichos y verdaderos) de tipo binario como *Cross-Entropy*; para intentar superar el

TABLA I  
ARQUITECTURA DE REDES.

Red	Tipo de capa	Tamaño de salida	Ajuste adicional
1	Entrada (x1)	(None, 6250, 22)	-
	LSTM 1	(None,6250,64)	Regularizer L2(0.001)
	Dropout 1	(None,6250,64)	0.2
	LSTM 2	(None,32)	Regularizer L2(0.001)
	Dropout 2	(None,32)	0.2
	Dense	(None,1)	Sigmoid
2	Red 1	(None,6250,22)	-
	LSTM 3	(None,32)	Regularizer L2(0.001)
	Dropout 3	(None,32)	0.2
3 <sup>a</sup>	Dense	(None,1)	Sigmoid
	Entradas (x22)	(None,6250,1)	-
	LSTM 1-22	(None,6250,1)	Regularizer L2(0.001)
	Concatenate	(None,6250,22)	-
	LSTM 23	(None,6250,128)	Regularizer L2(0.001)
	Dropout 1	(None,6250,128)	-
4	LSTM 24	(None,6250,64)	-
	Dropout 2	(None,6250,64)	-
	Dense	(None,1)	Sigmoid
	Red 3	(None,6250,22)	-
	Conv1D	(None,6250,128)	Kernel size=3 Pool size=2
	MaxPooling1D	(None,3125,128)	-
5 <sup>b</sup>	LSTM 1	(None,6250,128)	Regularizer L2(0.001)
	LSTM 2	(None,64)	Regularizer L2(0.001)
	Dense	(None,1)	Sigmoid
	Entrada (x1)	(None,6250,22)	-
	Conv1D 1-3 (3)	(None,6250,32)	Kernel=2,4,8; Strides=2
	Concatenate 1	(None,3125,96)	-
5 <sup>b</sup>	Conv1D 4-6	(None,1563,32)	Kernel=2,4,8; Strides=2
	Concatenate 2	(None,1563,96)	-
	Conv1D 7-9	(None,782,32)	Kernel=2,4,8; Strides=2
	Concatenate 3	(None,782,96)	-
	GRU 1	(None,782,32)	-
	GRU 2	(None,782,32)	-
	Concatenate 4	(None,782,64)	-
	GRU 3	(None,782,32)	-
	Concatenate 5	(None,782,64)	-
	GRU 4	(None,782,32)	-
	Concatenate 6	(None,782,96)	-
	GRU 5	(None,32)	-
Dense	(None,1)	Sigmoid	

<sup>a</sup> Basado en [28], [37]. <sup>b</sup> Basado en la red ChronoNet [32].

Se mantuvieron los términos en inglés utilizados con TensorFlow. LSTM: Long Short-Term Memory; Conv1D: capa convolucional unidimensional; GRU: Gated Recurrent Units; Concatenate: capa de concatenación para la salida de otras capas; Dropout: capa para eliminación de entradas aleatorias; Dense: capa densamente conectada; MaxPooling1D: Operación *max pooling*; Kernel o kernel size: tamaño de ventana de convolución; Strides: pasos de convolución; Regularizer L2: uso de penalidades en los parámetros de capa durante el entrenamiento del modelo; Sigmoid: Función sigmoidea para clasificación binaria.

problema de sobreajuste y forzar al sistema a aprender características de la señal, múltiples capas *Dropout* fueron usadas entre las capas de red [38]; un algoritmo de optimización de tipo Adam con una tasa de aprendizaje de  $1 \times 10^{-3}$ ; un máximo de 20 pasadas (*epochs*) y un punto de corte igual a 0.5 para la predicción de clases positivas (señal convulsiva). Además, usando un ajuste de parada temprana, se interrumpió el entrenamiento tan pronto la exactitud de validación no mejoró. Las curvas ROC y matrices de confusión obtenidas durante las evaluaciones mixtas y específicas se detallan en las Figs. 5 y 6 respectivamente. Todos los códigos implementados y datos de entrenamiento preprocesados, se encuentran disponibles en Code Ocean <https://codeocean.com/capsule/6306977/tree>.

El rápido sobreajuste de los datos durante el entrenamiento puede explicarse por dos razones: Problema intrínseco en las redes LSTM y los pocos datos de entrenamiento utilizados a pesar de la diversidad de las sesiones EEG. Obteniendo un mejor comportamiento por las redes simples (1 y 2). Sobre todo, el rendimiento alcanzado por la red 2 durante la evaluación mixta, a pesar de los bajos valores AUC (Fig. 5b). Pudiendo rectificarse al utilizar mayor cantidad de datos.

Por otro lado, considerando la importancia de la especificidad a la hora de minimizar el error de predicciones para pacientes sanos, las demás redes funcionan mejor. Así, las redes 4 y 5 sacrifican valores recall para aumentar una correcta

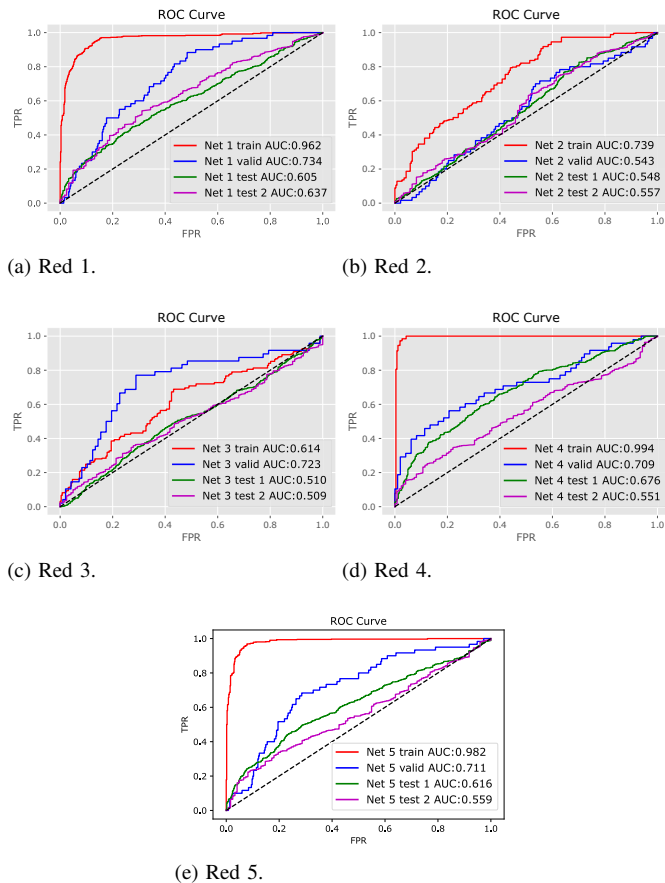


Fig. 5. Curvas ROC para los datos de entrenamiento, validación, evaluaciones mixta y específica.

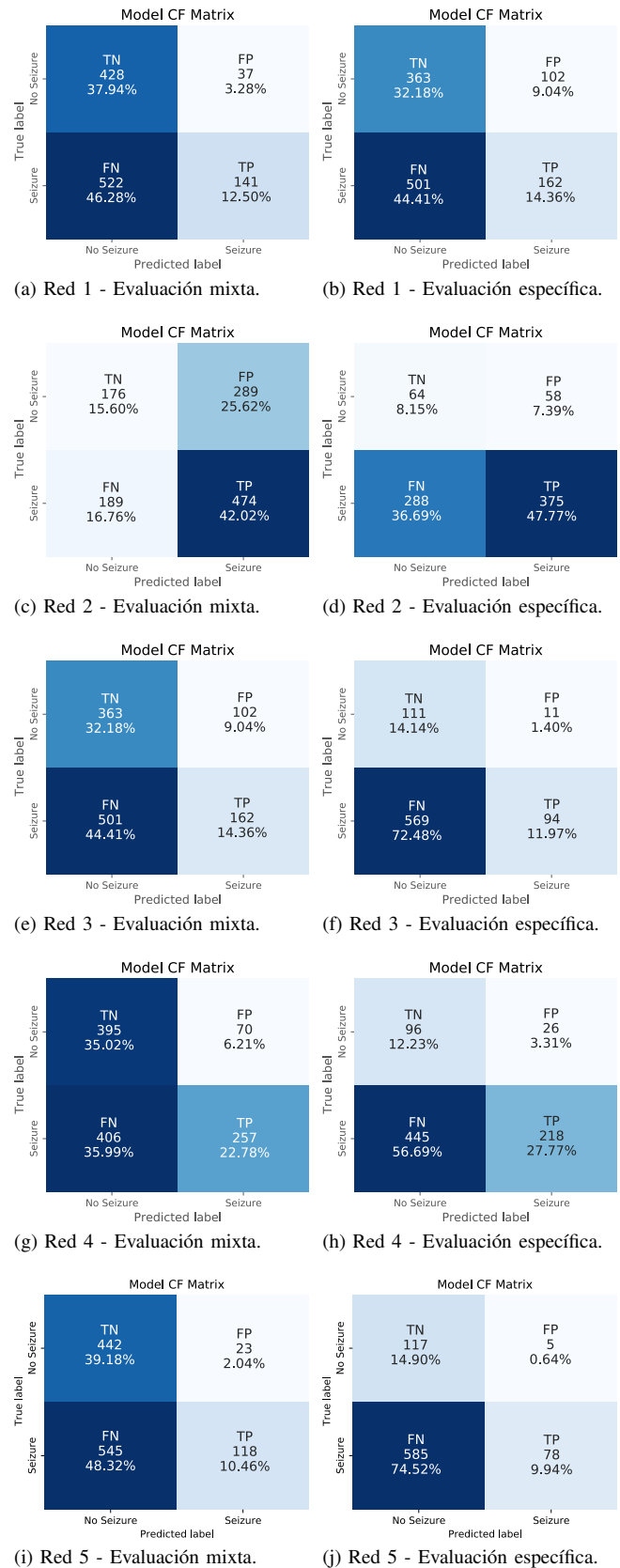


Fig. 6. Matrices de confusión para los distintos modelos.

TABLA II  
MÉTRICAS DE RENDIMIENTO PARA LOS GRUPOS DE ENTRENAMIENTO Y VALIDACIÓN.

Red	Exactitud	Recall	Precisión	Especificidad
Entrenamiento 1 <sup>a</sup>	85.6	87.1	84.5	84.0
Validación 1	82.2	48.3	39.2	87.8
Entrenamiento 2 <sup>b</sup>	67.3	93.0	61.4	41.6
Validación 2	50.2	61.7	16.3	48.4
Entrenamiento 3 <sup>b</sup>	58.8	54.9	59.5	62.6
Validación 3	71.0	52.1	56.8	80.4
Entrenamiento 4 <sup>b</sup>	97.5	1	95.2	94.9
Validación 4	73.8	43.8	65.6	88.7
Entrenamiento 5 <sup>a</sup>	92.8	89.6	95.7	95.9
Validación 5	82.2	11.7	23.3	93.8

<sup>a</sup> Datos con grupo grande. <sup>b</sup> Con grupo reducido.

TABLA III  
MÉTRICAS DE RENDIMIENTO PARA EL GRUPO DE EVALUACIÓN MIXTA: PACIENTE VS CONTROL.

Red	Exactitud	Recall	Precisión	Especificidad	F1-score
1	49.60	25.20	69.90	84.50	0.37
2	57.60	71.50	62.10	37.80	0.66
3	46.50	24.40	61.40	78.10	0.35
4	57.80	38.80	78.60	84.90	0.52
5	49.60	17.80	83.70	95.10	0.29

identificación de ausencia de convulsiones, que, desde el punto de vista clínico, es preferible en términos de costo-beneficio para una prueba diagnóstica. La clasificación errónea de pacientes solo genera gastos innecesarios, reclamos por error diagnóstico, etc. Llegando a superar el rendimiento humano en tareas similares con 61.1 % de especificidad [39]. Estas diferencias entre modelos podrían responderse al incremento de capas recurrentes (red 2 es superior a red 1), el uso de redes híbridas para mejorar la extracción de características (red 4 es superior a red 3) o la correcta combinación de ambas (red 5).

En cuanto a otros autores (Tabla V), las redes 2, 4 y 5 alcanzaron resultados prometedores. Así, la red 2 supera en 37.80 % de sensibilidad al estado del arte basado en la misma base de datos, para una detección de convulsiones entre pacientes y controles. Por su parte, la red 5 también mejora la especificidad en un 25.18 % para detecciones paciente-control, y hasta un 24.38 % usando exclusivamente datos de pacientes.

Por el contrario, el rendimiento obtenido con otras bases de datos provenientes de la University of Bonn (señales con un solo canal y sin artefactos) [23], [25]–[28], [34], Epilepsy Center of the University Hospital of Freiburg [33] y Children’s

TABLA IV  
MÉTRICAS DE RENDIMIENTO PARA EL GRUPO DE EVALUACIÓN ESPECÍFICA: PACIENTE (CONVULSIÓN) VS PACIENTE (NO CONVULSIÓN).

Red	Exactitud	Recall	Precisión	Especificidad	F1-score
1	28.40	16.60	92.40	92.60	0.28
2	55.90	56.60	86.600	52.50	0.87
3	26.10	14.20	89.5	91.0	0.24
4	40.0	32.90	89.30	78.70	0.48
5	24.80	11.80	94.0	95.90	0.21

TABLA V  
COMPARACIÓN DE NUESTRAS REDES Y EL ESTADO DEL ARTE.

Redes	Exactitud		Evaluación mixta/específica		
	Entren.	Valid.	Recall	Especificidad	AUC
LSTM [25] <sup>a</sup>	N/A	N/A	33.71	70.72	N/A
1D-CNN-RNN [31] <sup>b</sup>	99.16	82.27	N/A	N/A	N/A
Nuestra red 2	67.30	50.20	71.50/56.60	37.80/52.50	0.55/0.56
Nuestra red 4	97.50	73.80	38.80/32.90	84.90/78.70	0.67/0.55
Nuestra red 5	92.80	82.20	17.80/11.80	95.10/95.90	0.62/0.60
ChronoNet [32] <sup>b</sup>	90.60	86.57	N/A	N/A	N/A
1D-CNN + LSTM [40] <sup>a</sup>	N/A	N/A	N/A	N/A	0.80

<sup>a</sup> Usando TUSZ. <sup>b</sup> TUH Abnormal EEG Corpus.

Hospital of Boston-Massachusetts Institute of Technology [24], [30], presentan mejores métricas para la detección de convulsiones. Esto debido a la combinación de factores como (1) los altos valores de sensibilidad y bajos falsos positivos demandados por la comunidad de neurociencia [25] y, (2) la dificultad para entrenar modelos usando TUSZ. Por ejemplo, la exactitud obtenida por la red ChronoNet original es mayor a la red 5 debido a su diseño inicial para la detección de señales anormales, tarea más fácil que la detección de convulsiones.

Hay muchas limitaciones en este estudio. Entre ellas, el reducido número de sesiones consideradas para el entrenamiento supervisado de las redes, debido a los recursos limitados de computación en la nube con la plataforma Google Colab para aplicaciones en ML. Lo que explica el sobreajuste de los datos.

## V. CONCLUSIONES

La detección automática de convulsiones es una tarea crucial en el diagnóstico de epilepsia que podría superar las desventajas del diagnóstico visual, al servir como herramienta de asistencia al médico neurólogo y disminuir la dependencia de profesionales entrenados para una interpretación manual y prolongada de las señales. Además, resulta económicamente atractivo como herramienta clínica al reducir el tiempo diagnóstico, de monitoreo y planificar tratamientos oportunos que mejoren la calidad de vida del paciente. Pudiendo ser particularmente útil y necesaria en países en vías de desarrollo como la región Latinoamérica. En este trabajo, el uso de redes neuronales con arquitecturas simples y complejas, basadas principalmente en LSTM, demuestra ser flexible para la detección robusta de convulsiones mediante el manejo de señales EEG como series de tiempo, sin un extenso y complejo preprocesamiento inicial de datos crudos del Temple University Hospital (TUH) EEG Seizure Corpus.

A pesar de las limitaciones computacionales para procesar una gran cantidad de muestras y un diseño profundo de redes neuronales, una red simple LSTM, otra híbrida (capas LSTM y 1D-CNN) y la red ChronoNet (capas 1D-CNN, LSTM y GRU) muestran métricas prometedoras para una clasificación más robusta y binaria de convulsiones con sensibilidad y especificidad de hasta 71.50 % y 83.70 %, para una detección entre pacientes y controles; y hasta un 56.60 % y 95.90 %, para una detección específica de pacientes respectivamente. El siguiente paso será mejorar la arquitectura de redes y aumentar el número de datos, mejorando la extracción de características

y preprocesamiento de la señal, a fin de validar los modelos con datos locales que midan su verdadero impacto clínico.

#### REFERENCIAS

- [1] W. H. Organization, "Epilepsy." [Online]. Available at <https://www.who.int/news-room/fact-sheets/detail/epilepsy>, 2022.
- [2] R. S. Fisher, W. v. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel, "Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," *Epilepsia*, vol. 46, pp. 470–472, Apr. 2005.
- [3] C. E. Elger and C. Hoppe, "Diagnostic challenges in epilepsy: Seizure under-reporting and seizure detection," *The Lancet Neurology*, vol. 17, pp. 279–288, Mar. 2018.
- [4] W. H. Organization, *Epilepsy: A Public Health Imperative*. No. WHO/MSD/MER/19.2, World Health Organization, 2019.
- [5] M. Bosak, A. Słowik, R. Kacorzyk, and W. Turaj, "Implementation of the new ILAE classification of epilepsies into clinical practice — A cohort study," *Epilepsy & Behavior*, vol. 96, pp. 28–32, July 2019.
- [6] J. Cloyd, W. Hauser, A. Towne, R. Ramsay, R. Mattson, F. Gilliam, and T. Walczak, "Epidemiological and medical aspects of epilepsy in the elderly," *Epilepsy Research*, vol. 68, pp. 39–48, Jan. 2006.
- [7] J. Falco-Walter, "Epilepsy—Definition, Classification, Pathophysiology, and Epidemiology," *Seminars in Neurology*, vol. 40, pp. 617–623, Dec. 2020.
- [8] O. Devinsky, T. Spruill, D. Thurman, and D. Friedman, "Recognizing and preventing epilepsy-related mortality: A call for action," *Neurology*, vol. 86, pp. 779–786, Feb. 2016.
- [9] A. Tanaka, N. Akamatsu, T. Shouzaki, T. Toyota, M. Yamano, M. Nakagawa, and S. Tsuji, "Clinical characteristics and treatment responses in new-onset epilepsy in the elderly," *Seizure*, vol. 22, pp. 772–775, Nov. 2013.
- [10] W. A. Hauser, "An unparalleled assessment of the global burden of epilepsy," *The Lancet Neurology*, vol. 18, pp. 322–324, Apr. 2019.
- [11] S. Ammar and M. Senouci, "Seizure detection with single-channel EEG using Extreme Learning Machine," in *2016 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA)*, (Sousse, Tunisia), pp. 776–779, IEEE, Dec. 2016.
- [12] P. Thodoroff, J. Pineau, and A. Lim, "Learning Robust Features using Deep Learning for Automatic Seizure Detection," 2016.
- [13] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *Journal of Neural Engineering*, vol. 16, p. 031001, June 2019.
- [14] S. Ahufinger, P. Balugo, M. M. González, E. Pequeño, H. González, and P. Herrero, "A User-centered Smartphone Application for Wireless EEG and its Role in Epilepsy," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 5, no. 6, p. 43, 2019.
- [15] D. S. W. Ting, L. Carin, V. Dzau, and T. Y. Wong, "Digital technology and COVID-19," *Nature Medicine*, vol. 26, pp. 459–461, Apr. 2020.
- [16] S. Hurtado, J. García-Nieto, A. Popov, and I. Navas-Delgado, "Human Activity Recognition From Sensorised Patient's Data in Healthcare: A Streaming Deep Learning-Based Approach," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. In Press, no. In Press, p. 1, 2022.
- [17] S. Afzal, M. Maqsood, U. Khan, I. Mehmood, H. Nawaz, F. Aadil, O.-Y. Song, and Y. Nam, "Alzheimer Disease Detection Techniques and Methods: A Review," *International Journal of Interactive Multimedia and Artificial Intelligence*, vol. 6, no. 7, p. 26, 2021.
- [18] S. Kiranyaz, T. Ince, M. Zabihi, and D. Ince, "Automated patient-specific classification of long-term Electroencephalography," *Journal of Biomedical Informatics*, vol. 49, pp. 16–31, June 2014.
- [19] A. Shoeibi, M. Khodatars, N. Ghassemi, M. Jafari, P. Moridian, R. Alizadehsani, M. Panahiazar, F. Khozeimeh, A. Zare, H. Hosseini-Nejad, A. Khosravi, A. F. Atiya, D. Aminshahidi, S. Hussain, M. Rouhani, S. Nahavandi, and U. R. Acharya, "Epileptic Seizures Detection Using Deep Learning Techniques: A Review," *International Journal of Environmental Research and Public Health*, vol. 18, p. 5780, May 2021.
- [20] B. Abbasi and D. M. Goldenholz, "Machine learning applications in epilepsy," *Epilepsia*, vol. 60, pp. 2037–2047, Oct. 2019.
- [21] D. O. Nahmias, E. F. Civillico, and K. L. Kontson, "Deep learning and feature based medication classifications from EEG in a large clinical data set," *Scientific Reports*, vol. 10, p. 14206, Dec. 2020.
- [22] K. P. Thanaraj, B. Parvathavarthini, U. J. Tanik, V. Rajinikanth, S. Kadry, and K. Kamalanand, "Implementation of Deep Neural Networks to Classify EEG Signals using Gramian Angular Summation Field for Epilepsy Diagnosis," 2020.
- [23] I. Ullah, M. Hussain, E.-u.-H. Qazi, and H. Aboalsamh, "An automated system for epilepsy detection using EEG brain signals based on deep learning approach," *Expert Systems with Applications*, vol. 107, pp. 61–71, Oct. 2018.
- [24] L. Vidyaratne, A. Glandon, M. Alam, and K. M. Iftekharuddin, "Deep recurrent neural network for seizure detection," in *2016 International Joint Conference on Neural Networks (IJCNN)*, (Vancouver, BC, Canada), pp. 1202–1207, IEEE, July 2016.
- [25] M. Golmohammadi, S. Ziyabari, V. Shah, S. L. de Diego, I. Obeid, and J. Picone, "Deep Architectures for Automated Seizure Detection in Scalp EEGs," 2017.
- [26] S. S. Talathi, "Deep Recurrent Neural Networks for seizure detection and early seizure detection systems," 2017.
- [27] X. Chen, J. Ji, T. Ji, and P. Li, "Cost-Sensitive Deep Active Learning for Epileptic Seizure Detection," in *Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics*, (Washington DC USA), pp. 226–235, ACM, Aug. 2018.
- [28] R. Hussein, H. Palangi, R. Ward, and Z. J. Wang, "Epileptic Seizure Detection: A Deep Learning Approach," 2018.
- [29] K. Fukumori, H. T. Thu Nguyen, N. Yoshida, and T. Tanaka, "Fully Data-driven Convolutional Filters with Deep Learning Models for Epileptic Spike Detection," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Brighton, United Kingdom), pp. 2772–2776, IEEE, May 2019.
- [30] X. Yao, Q. Cheng, and G.-Q. Zhang, "Automated Classification of Seizures against Nonseizures: A Deep Learning Approach," 2019.
- [31] S. Roy, I. Kiral-Kornek, and S. Harrer, "Deep Learning Enabled Automatic Abnormal EEG Identification," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Honolulu, HI), pp. 2756–2759, IEEE, July 2018.
- [32] S. Roy, I. Kiral-Kornek, and S. Harrer, "ChronoNet: A Deep Recurrent Neural Network for Abnormal EEG Identification," 2018.
- [33] M. Geng, W. Zhou, G. Liu, C. Li, and Y. Zhang, "Epileptic Seizure Detection Based on Stockwell Transform and Bidirectional Long Short-Term Memory," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, pp. 573–580, Mar. 2020.
- [34] A. Verma and R. R. Janghel, "Epileptic Seizure Detection Using Deep Recurrent Neural Networks in EEG Signals," in *Advances in Biomedical Engineering and Technology*, pp. 189–198, Singapore: Springer Singapore, 2021.
- [35] V. Shah, E. von Weltin, S. Lopez, J. R. McHugh, L. Veloso, M. Golmohammadi, I. Obeid, and J. Picone, "The Temple University Hospital Seizure Detection Corpus," *Frontiers in Neuroinformatics*, vol. 12, p. 83, Nov. 2018.
- [36] I. Obeid and J. Picone, "The Temple University Hospital EEG Data Corpus," *Frontiers in Neuroscience*, vol. 10, May 2016.
- [37] C. Spampinato, S. Palazzo, I. Kavasidis, D. Giordano, N. Souly, and M. Shah, "Deep Learning Human Mind for Automated Visual Classification," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (Honolulu, HI), pp. 4503–4511, IEEE, July 2017.
- [38] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, no. 56, pp. 1929–1958, 2014.
- [39] C. B. Swisher, C. R. White, B. E. Mace, K. E. Dombrowski, A. M. Husain, B. J. Kolls, R. R. Radtke, T. T. Tran, and S. R. Sinha, "Diagnostic Accuracy of Electrographic Seizure Detection by Neurophysiologists and Non-Neurophysiologists in the Adult ICU Using a Panel of Quantitative EEG Trends," *Journal of Clinical Neurophysiology*, vol. 32, pp. 324–330, Aug. 2015.
- [40] K. Lee, H. Jeong, S. Kim, D. Yang, H.-C. Kang, and E. Choi, "Real-Time Seizure Detection using EEG: A Comprehensive Comparison of Recent Approaches under a Realistic Setting," Mar. 2022.



**Sergio Augusto Jacobo Zavaleta** , received his Bachelor's Degree in Mechatronics Engineering (2019) from the Universidad Nacional de Trujillo, Peru. His research interests include design of clinical robots and biomedical applications based on Artificial Intelligence using Machine Learning.



**Jorge Juan Zavaleta Gavidia** , received the Ph.D in Systems and Computer Engineering Federal University of Rio de Janeiro (UFRJ) in 2017. He received the title of Master in Computer Science from the Federal University of Rio Grande do Sul (UFRGS) in 1997. He received the title of Licentiate in Mathematics from the Universidad Nacional de Trujillo (UNT) in 1998 and received a bachelor's degree in Physical and Mathematical Sciences from the UNT in 1992. Professor of Computing and currently a researcher of postdoctoral at the State University of

Rio de Janeiro (UERJ) in the project CAPES-Telemedecine and Medical Data Analysis. He is interested in topics research related to Data Science, Artificial Intelligence and Machine Learning.