

ANÁLISIS DE IMÁGENES RAYOS X CEREBRALES USANDO REDES NEURONALES CONVOLUCIONALES VGGNET

BRAIN X-RAY IMAGE ANALYSIS USING VGGNET CONVOLUTIONAL NEURAL NETWORKS

PhD. Angel Cuenca¹, Ing. Edgar Basurto², Ing. Rafael Larrea³

Resumen

El presente estudio aborda el desarrollo y entrenamiento de un modelo de detección de tumores cerebrales utilizando redes neuronales convolucionales (CNN) basadas en VGGNet. La investigación se centra en la aplicación de esta tecnología de inteligencia artificial en la interpretación de imágenes médicas de radiografías cerebrales para la detección temprana de patologías. Se inicia con la recopilación de un conjunto de datos de imágenes de radiografías cerebrales, que se divide en dos categorías: "Tumor" y "Healthy". Se aplican técnicas de preprocesamiento para normalizar las imágenes. Se elige la arquitectura VGGNet16 como base para el modelo CNN, se adaptó la capa de salida para ser usada una clasificación binaria. Durante el entrenamiento, se implementa un Learning Rate Schedule y EarlyStopping para evitar el sobreaprendizaje, se obtiene una pérdida del 3.92% y una exactitud del 98.68%. En conclusión, la investigación demuestra que la combinación de la arquitectura VGGNet16 y técnicas de preprocesamiento con ajustes cuidadosos de hiperparámetros resulta en un modelo de CNN altamente efectivo para la detección de tumores cerebrales en imágenes de radiografías. Sin embargo, se destaca la importancia de considerar limitaciones como el tamaño del conjunto de datos, la variabilidad de las imágenes y la necesidad de validación clínica antes de aplicar el modelo en un entorno médico real.

Palabras clave: Detección temprana, Patologías, Radiografías cerebrales, Redes neuronales convolucionales (CNN), Tumores cerebrales, VGGNet.

Abstract

The present study addresses the development and training of a brain tumor detection model using convolutional neural networks (CNN) based on VGGNet. The research focuses on the application of this artificial intelligence technology in the interpretation of medical images of brain X-rays for the early detection of pathologies. It starts with collecting a data set of brain X-ray images, which is divided into two categories: "Tumor" and "Healthy". Preprocessing techniques are applied to normalize the images. The VGGNet16 architecture is chosen as the basis for the CNN model, the output layer was adapted to use a binary classification. During training, a Learning Rate Schedule and EarlyStopping are implemented to avoid overlearning, a loss of 3.92% and an accuracy of 98.68% are obtained. In conclusion, the research demonstrates that the combination of the VGGNet16 architecture and preprocessing techniques with careful hyperparameter tuning results in a highly effective CNN model for the detection of brain tumors in X-ray images. However, it is important to consider limitations such as the size of the data set, the variability of the images, and the need for clinical validation before applying the model in a real medical setting.

Keywords: Early detection, Pathologies, Brain radiographs, Convolutional neural networks (CNN), Brain tumors, VGGNet.

¹Facultad de Ciencias Matemáticas y Físicas, Universidad de Guayaquil - Ecuador.

²Investigador Independiente - edabaro0191@gmail.com.

³Investigador Independiente - rlarrea@gmail.com

1. Introducción

El análisis de imágenes médicas ha experimentado un notable avance gracias a los desarrollos en inteligencia artificial y aprendizaje profundo. Sin embargo, su interpretación precisa puede ser un desafío debido a la complejidad anatómica y a las sutilezas presentes en las patologías cerebrales [1].

En este contexto, las redes neuronales convolucionales han demostrado un gran potencial para mejorar la precisión y eficiencia del diagnóstico radiológico. En particular, las CNN¹ se han convertido en una herramienta poderosa para el análisis de imágenes [2], entre ellas se destaca VGGNet² por su efectividad.

En sí las CNN es un algoritmo de aprendizaje profundo diseñados específicamente para el procesamiento de imágenes. Utilizan capas de convolución para extraer características relevantes, permitiéndoles aprender patrones y estructuras complejas presentes en los datos. La arquitectura VGGNet, desarrollada por la Universidad de Oxford, se destaca por su simplicidad y profundidad [3]. Compuesta por una serie de capas le da la capacidad para aprender representaciones jerárquicas de las imágenes [3].

El preprocesamiento de las imágenes cerebrales es un aspecto importante antes de su entrada en la red neuronal. Ya que requiere de normalizar las imágenes, eliminar artefactos y resaltar características relevantes [4]. Además, se requiere de un alto conjunto de datos para un entrenamiento efectivo y que mejore la generalización del modelo.

En el documento se plantea desarrollar un sistema de análisis de imágenes de radiografías cerebrales utilizando la arquitectura de redes neuronales convolucionales VGGNet, para ello se requiere comprender los fundamentos de las redes neuronales convolucionales y la arquitectura VGGNet, incluyendo su estructura, funcionamiento y capacidades; se propone caso de estudio: diseño, entrenamiento y análisis de una red neuronal convolucional VGGNet, cerrando el documento con las conclusiones del caso.

2. Materiales y Métodos

2.1. Inteligencia Artificial

La inteligencia artificial (IA) se refiere a la simulación de procesos de inteligencia humana mediante la programación de sistemas informáticos o máquinas [5]. Estos sistemas están diseñados para realizar tareas que normalmente requieren inteligencia humana, como

el razonamiento, el aprendizaje, la resolución de problemas, la percepción visual, el reconocimiento de voz, la toma de decisiones y la comprensión del lenguaje natural.

La IA abarca un amplio espectro de enfoques y técnicas, desde algoritmos simples hasta sistemas más complejos basados en redes neuronales y aprendizaje profundo. [6] Algunas áreas de la IA incluyen:

- **Aprendizaje automático (Machine Learning):** Implica el desarrollo de algoritmos y modelos que permiten a las máquinas mejorar su rendimiento en una tarea específica a medida que se les proporcionan más datos y retroalimentación.
- **Redes Neuronales Artificiales:** Modelos inspirados en la estructura y funcionamiento del cerebro humano que se utilizan en aprendizaje profundo para resolver problemas complejos, como reconocimiento de imágenes y procesamiento de lenguaje natural.
- **Procesamiento del Lenguaje Natural (NLP):** Se centra en permitir que las máquinas comprendan, interpreten y generen lenguaje humano de manera efectiva. Esto incluye tareas como traducción automática, generación de texto y análisis de sentimientos.
- **Visión por Computadora:** Implica enseñar a las máquinas a interpretar y comprender imágenes y videos, permitiéndoles identificar objetos, rostros y patrones visuales.
- **Robótica:** Se refiere al diseño y desarrollo de robots que pueden realizar tareas físicas y cognitivas, interactuando con el entorno de manera autónoma o asistida.
- **Aprendizaje Reforzado (Reinforcement Learning):** Implica entrenar a las máquinas para tomar decisiones mediante la interacción con su entorno y la retroalimentación recibida en función de las acciones tomadas.
- **Inteligencia Artificial General (AGI):** Representa la idea de una máquina con capacidad de inteligencia similar a la humana, capaz de realizar una amplia gama de tareas intelectuales y aprender de manera independiente en diversos dominios.

La inteligencia artificial tiene aplicaciones en diversas industrias, como la atención médica, la automoción, la financiación, la manufactura, la educación y muchos otros campos. Su desarrollo y avance continúan

¹CNN - Convolutional Neural Network trad. *Redes Neuronales Convolucionales*

²VGGNet – Visual Geometry Group Network

transformando la forma en que interactuamos con la tecnología y cómo se abordan los desafíos en diferentes áreas de la vida. [7]

2.2. CNN - Convolutional Neural Network

Una Red Neuronal Convolutiva (CNN, por sus siglas en inglés, Convolutional Neural Network) es un tipo de arquitectura de redes neuronales profundas especialmente diseñada para procesar y analizar datos de tipo grid³, como imágenes y videos [8]. Las CNN se han vuelto fundamentales en el campo de la visión por computadora y han revolucionado la forma en que se abordan tareas como el reconocimiento de objetos, la detección de características y la segmentación de imágenes [9]. Las características distintivas de una CNN incluyen:

- **Convoluciones:** Las convoluciones son operaciones matemáticas que aplican un filtro (kernel) a una región de la imagen para extraer características relevantes, como bordes, texturas y patrones. Estas operaciones se realizan en varias capas de la red para capturar características de diferentes niveles de abstracción.
- **Capas de Pooling:** Estas capas reducen el tamaño de la representación espacial de las características, disminuyendo así la cantidad de información y los parámetros de la red. La operación común de pooling es el max-pooling, que selecciona el valor máximo de un conjunto de valores en una región.
- **Capas de Activación:** Las funciones de activación, como la función ReLU (Rectified Linear Unit), introducen no linealidad en la red y permiten que la CNN capture relaciones y patrones más complejos en los datos.
- **Capas Totalmente Conectadas:** Al final de la red, se pueden utilizar capas totalmente conectadas para realizar tareas de clasificación u otras tareas específicas. Estas capas combinan las características extraídas de las capas anteriores y generan la salida final.
- **Aprendizaje Jerárquico de Características:** A través de sus capas, las CNN aprenden automáticamente jerarquías de características, desde características simples (como bordes) en las capas iniciales hasta características más abstractas y complejas (como formas y objetos) en las capas posteriores.

Las CNN se han utilizado con éxito en una variedad de aplicaciones, como:

- Reconocimiento de objetos y detección de objetos en imágenes y videos.
- Clasificación de imágenes en categorías específicas (por ejemplo, identificar diferentes especies de animales).
- Segmentación semántica, donde se asigna una etiqueta a cada píxel de una imagen.
- Generación de imágenes y estilo de imágenes.
- Procesamiento de imágenes médicas y análisis de radiografías.
- Conducción autónoma y sistemas de detección de peatones en vehículos.

Las CNN han demostrado ser altamente eficaces en la extracción de características y la representación de datos visuales, lo que ha llevado a avances significativos en la visión por computadora y en aplicaciones relacionadas con el análisis de imágenes y videos [10].

2.3. VGG - Visual Geometry Group Network

La Red Visual Geometry Group (VGG) es una arquitectura de red neuronal convolutiva (CNN) que se destacó en el Concurso de Reconocimiento Visual de ImagenNet en 2014 [11]. Fue desarrollada por el Grupo de Geometría Visual (Visual Geometry Group) en la Universidad de Oxford. La arquitectura VGG se caracteriza por su profundidad y simplicidad en la estructura de capas, lo que la hace fácil de entender y entrenar [12]. Las principales características de la arquitectura VGG son:

- **Profundidad:** A diferencia de algunas arquitecturas más tempranas que utilizaban un número limitado de capas, VGG emplea una estructura profunda con múltiples capas convolucionales y capas totalmente conectadas. Su enfoque en la profundidad contribuye a su capacidad para aprender representaciones más ricas y complejas de los datos.
- **Tamaño de Filtro Pequeño:** VGG utiliza filtros convolucionales de tamaño pequeño (3x3) en todas sus capas convolucionales, seguidos de una función de activación ReLU⁴. Esta elección de filtros pequeños permite la captura de características más locales y facilita el aprendizaje de representaciones detalladas de la imagen.
- **Max-Pooling:** Después de cada conjunto de capas convolucionales, VGG utiliza capas de max-pooling para reducir el tamaño espacial de las características y disminuir la cantidad de parámetros en la red.

³Datos tipos grid, refiere a datos presentados en una matriz

⁴ReLU (Rectified Linear Unit), función de activación no lineal que se usa en redes neuronales convolucionales para introducir no linealidad y eliminar valores negativos.

- **Capas Totalmente Conectadas:** Al final de la red, VGG utiliza capas totalmente conectadas para llevar a cabo tareas específicas, como la clasificación. Estas capas combinan las características extraídas en las capas anteriores y generan la salida final.
- **Variantes:** VGG se presenta en varias variantes según la profundidad, como VGG16 y VGG19, que indican el número de capas en la red.

Aunque VGG fue una de las arquitecturas clave en la evolución de las redes neuronales convolucionales, en la práctica, las arquitecturas más modernas, como las redes ResNet⁵, Inception⁶ y las basadas en atención, han superado sus resultados en términos de precisión y eficiencia en diversas tareas de visión por computadora [13]. Sin embargo, VGG sigue siendo relevante para entender los principios básicos de las CNN y su importancia en el desarrollo de arquitecturas más avanzadas.

2.4. Imágenes de Rayos X

Las imágenes de rayos X son un tipo de técnica de diagnóstico médico que utiliza radiación electromagnética de alta energía para crear imágenes del interior del cuerpo humano. Esta tecnología permite a los médicos visualizar estructuras internas como huesos, órganos y tejidos blandos sin necesidad de cirugía invasiva [14].

En una imagen de rayos X, los rayos X pasan a través del cuerpo y son absorbidos en diferentes grados por diferentes tejidos, lo que resulta en una imagen en blanco y negro que muestra contrastes entre las partes densas (como los huesos) y las partes menos densas (como los tejidos blandos) [15]. Las imágenes de rayos X son especialmente útiles para detectar fracturas óseas, evaluar el estado de los órganos internos, localizar objetos extraños en el cuerpo y ayudar en la planificación de procedimientos médicos.

Algunas aplicaciones comunes de las imágenes de rayos X incluyen:

- **Radiografías Torácicas:** Para examinar los pulmones y el corazón, y detectar condiciones como neumonía, tuberculosis o enfermedades cardíacas.
- **Radiografías Dentales:** Para evaluar la salud dental y diagnosticar problemas dentales y de la mandíbula.

- **Radiografías Óseas:** Para detectar fracturas, evaluar el crecimiento óseo en niños y diagnosticar condiciones como la osteoporosis.
- **Tomografía Computarizada (TC):** Una forma más avanzada de imagen de rayos X que crea imágenes tridimensionales detalladas del cuerpo, útiles para diagnósticos más precisos.
- **Angiografía:** Utilizada para visualizar los vasos sanguíneos y detectar bloqueos o anomalías en el flujo sanguíneo.
- **Mamografía:** Una técnica de rayos X específica para el examen de los senos, que se utiliza para detectar tempranamente el cáncer de mama.

Es importante destacar que, si bien las imágenes de rayos X son valiosas herramientas de diagnóstico, también implican la exposición a radiación ionizante. Los profesionales médicos toman precauciones para minimizar la exposición del paciente a la radiación y evalúan cuidadosamente los beneficios frente a los riesgos antes de realizar estudios de imágenes de rayos X [16].

2.5. Planteamiento

Se desea realizar el entrenamiento de un modelo CNN en VGGNet16 para lo cual se acudió a Kaggle⁷ con la finalidad de obtener una fuente de imágenes de RX cerebrales, se encontró el dataset de imágenes Brian Tumor Dataset⁸ con una clasificación de imágenes entre Brain Tumor (tumor) y Healthy (sanos); uso de COLAB⁹ como herramienta por el uso de cuadernos de Jupiter y lenguaje Python ideales para entrenamientos en inteligencia artificial.

2.6. Análisis de datos

Evalutando el dataset descargado, se dispone a seleccionar y mover al azar 58 imágenes de ambos segmentos los cuales serán usados para la evaluación del modelo, se crea archivo csv con data de respaldo identificando las imágenes con tumor y no-tumor, posterior se comprime los directorios y se hospeda en GitHub¹⁰ para uso público y del estudio. Detalle de directorios:

- **No_Brain_Tumor:** directorio que contiene banco de imágenes de rayos x de pacientes que no tienen tumor cerebral, contiene 2057 imágenes.
- **Yes_Brain_Tumor:** directorio que contiene banco de imágenes de rayos x de pacientes que

⁵ResNet, red neuronal profunda que usa conexiones residuales para mejorar el aprendizaje de las capas.

⁶Inception, red neuronal convolucional con 48 capas de profundidad.

⁷Kaggle, plataforma web que reúne dataset del todo el mundo.

⁸Brian Tumor Dataset, Kaggle link <https://www.kaggle.com/datasets/preetviradiya/brian-tumor-dataset>

⁹COLAB, desarrollador de programas de Python para escribir y ejecutar códigos en un navegador web.

¹⁰Link de archivos <https://github.com/Rafael1108/datasheet/tree/main/BrainTumorData>

presentan tumor cerebral, contiene 2048 imágenes.

- **Validate:** directorio que contiene banco de imágenes para evaluar el modelo, contiene 58 imágenes y 1 archivo csv.

2.6.1. Clasificación, segmentación y normalizado

En el cuaderno jupyter de COLAB se realiza la carga de las imágenes, con el soporte del paquete panda y la función `train_test_split` se establece la clasificación de las colecciones de para el entrenamiento, validación y pruebas:

```
*****
Total de Imágenes: 4542
Entrenar: 2271 | 50.0 % => [yes tumor, 1235] [no tumor, 1036]
Validar : 1135 | 24.9 % => [yes tumor, 621] [no tumor, 514]
Pruebas : 1136 | 25.1 % => [yes tumor, 629] [no tumor, 507]
*****
```

Figura 1. Estadística de datos a ser usados por el modelo.

Evaluando la metadata de las imágenes se puede apreciar que no disponen un estándar en la dimensión de las imágenes, se realiza el redimensionamiento de estas a 224 de width y 224 de height.

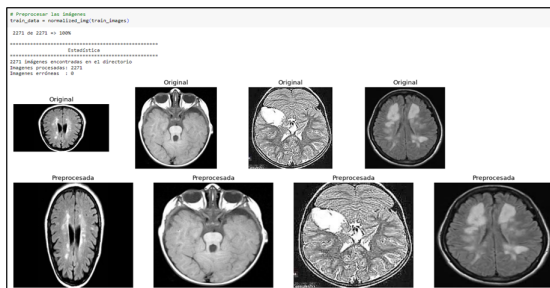


Figura 2. Resumen de procesamiento de imágenes por redimensionamiento.

2.7. Modelo

Dentro de los paquetes de tensorflow se puede importar la red del modelo para VGGNet16, sin embargo, hay que hacer los ajustes de entrada y salida dependiendo del tipo de clasificación que se use. A continuación, se muestra las librerías usadas para el entrenamiento del modelo.

```
#!pip install tensorflow
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.applications import VGG16
from tensorflow.keras.layers import Dense, GlobalAveragePooling2D
from tensorflow.keras.models import Model
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.callbacks import EarlyStopping
```

Figura 3. Paquetes o librerías usadas.

Para el estudio se plantea un caso de clasificación binaria por lo cual la última capa corresponde a `sigmoid` y al compilar el modelo se establece la perdida como `binary_crossentropy`.

```
# Cargar el modelo VGG16 sin las capas densas (fully connected)
base_model = VGG16(weights='imagenet', include_top=False, input_shape=(224, 224, 3))
```

Figura 4. Modelo VGG16.

```
[ ] # Agregar nuevas capas al modelo
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(128, activation='relu')(x)
predictions = Dense(1, activation='sigmoid')(x)
```

Figura 5. Nuevas capas al modelo, se establece la salida binaria.

```
[ ] #Compilar el modelo

model.compile(optimizer=optimizer, loss='binary_crossentropy', metrics=['accuracy'])
```

Figura 6. Modelo compilado para clasificación binaria.

2.8. Ajustes en el entrenamiento

Se estableció los siguientes parámetros:

- Learning Rate Schedule = 0.0001
- Epochs = 100
- EarlyStopping => (monitor='val_loss',
patience=5, verbose=1,
restore best weights=True)

2.8.1. Modelo VGGNet16 con clasificación binaria

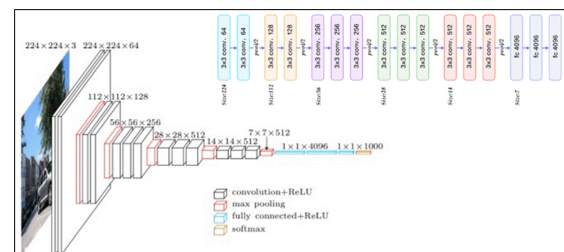


Figura 7. Arquitectura VGGNet 16, explicación gráfica.

ARQUITECTURA DEL MODELO		
Model: "model"		
Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
global_average_pooling2d (GlobalAveragePooling2D)	(None, 512)	0
dense (Dense)	(None, 128)	65664
dense_1 (Dense)	(None, 1)	129
Total params: 14,780,481		
Trainable params: 14,780,481		
Non-trainable params: 0		

Figura 8. Arquitectura VGGNet 16, vista de la arquitectura del modelo configurado.

3. Resultados y Discusión

Una vez entrenado el modelo con el marco establecido en el capítulo anterior, se obtiene corte en epoch 23 por tener activo **EarlyStopping** en la métrica de **val_loss**, con esto se evita un sobre aprendizaje del modelo; es decir, cuando el valor de la pérdida es constante y no sufre cambios consecutivos, el modelo deja de aprender porque llegó a su máximo en epoch 18 y **EarlyStopping** hace que se termine el proceso.

```

71/71 [=====] - 50s 816ms/step - loss: 0.0268 - accuracy: 0.9899 - val_loss: 0.0813 - val_accuracy: 0.9789
Epoch 21/100
71/71 [=====] - 51s 716ms/step - loss: 0.0476 - accuracy: 0.9841 - val_loss: 0.0790 - val_accuracy: 0.9718
Epoch 22/100
71/71 [=====] - 53s 740ms/step - loss: 0.0306 - accuracy: 0.9886 - val_loss: 0.0673 - val_accuracy: 0.9815
Epoch 23/100
71/71 [=====] - ETA: 0s - loss: 0.0367 - accuracy: 0.9886Restoring model weights from the end of the best epoch: 18.
71/71 [=====] - 53s 746ms/step - loss: 0.0367 - accuracy: 0.9886 - val_loss: 0.0555 - val_accuracy: 0.9797
Epoch 23: early stopping

```

Figura 9. Histórico de entrenamiento del modelo. [1]

3.1. Validación

En la evaluación del modelo con el segmento de imágenes para pruebas se pudo llegar a una pérdida de 3.92% y una exactitud del 98.68% siendo buenos rangos para la predicción.

A continuación, se muestra los histogramas para la pérdida y precisión:

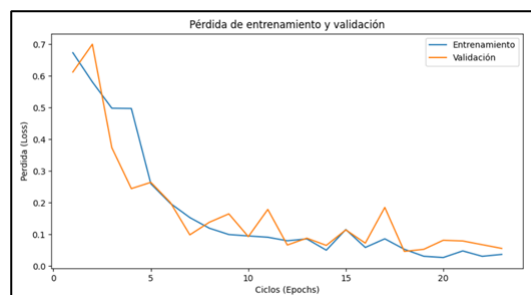


Figura 10. Histograma de pérdida. [1]

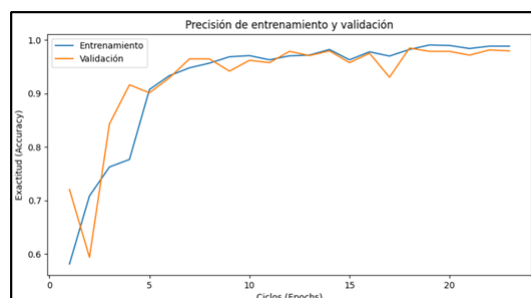


Figura 11. Histograma de precisión. [1]

3.2. Predicción

Culminada las fases anteriores se dispone de un modelo entrenado y dispuesto a dar predicciones con una exactitud del 98%, esto quiere decir que es capaz de identificar si existe o no tumor en una RX Cerebral. Para validar lo dicho se toma las 58 imágenes separadas al inicio del ejercicio, el modelo las recibe para predecir si tienen o no tumor, al final se hace una comparación con la matriz del archivo csv el cual contiene datos clasificatorios previos.

A continuación, se muestra gráfica de la curva ROC la cual nos muestra el resultado de la exactitud predicción realizada en estas 58 imágenes.

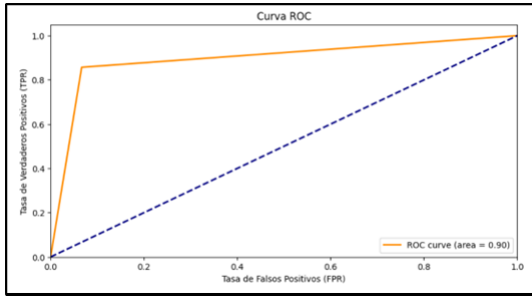


Figura 12. Histograma curva ROC. [1]

Según el estudio realizado en 2020 por Ejaz Ul Haq, Huang Jianjun, Kang Li, Hafeez Ul Haq & Tijiang Zhang ("An MRI-based deep learning approach for efficient classification of brain tumors") [17] el cual se centró en la detección y diagnóstico de tumores cerebrales utilizando escaneos de resonancia magnética (MRI¹¹) y redes neuronales convolucionales, los autores exploraron la capacidad de las redes neuronales para identificar patrones sutiles en las imágenes de MRI, los resultados experimentales demuestran una precisión del 97,3% y un coeficiente de similitud de dados (DSC¹²) del 95,8% en la tarea de clasificar el tumor cerebral como gliomas, meningiomas o tumores hipofisarios logrados por la primera arquitectura CNN propuesta, mientras que la segunda arquitectura CNN propuesta logró una precisión del 96,5% con un DSC del 94,3% en la tarea de clasificar los grados de glioma como HGG¹³ o LGG¹⁴.

Como se muestra en el estudio, basados en artículos científicos similares con alto grado de precisión, hemos planteado un caso con un segmento de imágenes de RX Cerebrales clasificados por tumorales y no tumorales; como VGGNet-16 se caracteriza por tener una arquitectura profunda y compleja, argumentada por 16 capas convolucionales y 3 capas totalmente conectadas. Estas capas facilitan a extraer características de las imágenes, mientras que las 3 capas son usadas para la clasificación. A lo largo de las etapas, se tomaron decisiones y ajustes clave que contribuyeron al éxito del modelo.

- Selección y Preprocesamiento de Datos: Se utilizó el dataset Brain Tumor Dataset, que contenía imágenes de radiografías cerebrales clasificadas en "Tumor" y "Healthy". Se llevó a cabo una selección aleatoria de imágenes para la evaluación y se creó un archivo CSV para respaldar esta selección. Las imágenes se redimensionaron a un tamaño uniforme de 224x224 píxeles para garantizar la entrada consistente al modelo.

- Arquitectura del Modelo VGGNet16: La arquitectura VGGNet16, conocida por su simplicidad y profundidad, se utilizó como base. La última capa de clasificación se adaptó para la clasificación binaria utilizando una función de activación sigmoid y la pérdida se estableció como binary_crossentropy. Se configuró un Learning Rate Schedule y se aplicó EarlyStopping para evitar el sobreaprendizaje.

- Entrenamiento y Evaluación del Modelo: El modelo se entrenó en un total de 100 epochs, pero se utilizó EarlyStopping para finalizar el proceso en el epoch 23 debido a que la métrica de val_loss dejó de mejorar. Se logró una pérdida del 3.92% y una exactitud del 98.68% en la evaluación del conjunto de pruebas, indicando un alto rendimiento en la predicción de imágenes cerebrales.

3.3. Análisis Comparativo

Tanto el presente estudio como el estudio de MRI lograron altos niveles de precisión en la predicción de tumores cerebrales, con un 98.68% y un 97.3% respectivamente. Esto indica que ambas metodologías y enfoques de entrenamiento son efectivos en la detección de patologías cerebrales.

Ambos estudios optaron por utilizar la arquitectura VGGNet16, lo que sugiere que esta arquitectura en particular ha demostrado ser eficaz en la clasificación de imágenes médicas. Sin embargo, el estudio de MRI pudo haber realizado ajustes específicos en la estructura de la red o en los hiperparámetros para optimizar aún más el rendimiento.

El estudio de MRI utilizó un conjunto de datos similar en tamaño al mostrado en el documento, con 3062 imágenes en total. Esto demuestra que un conjunto de datos suficientemente grande puede ser crucial para el éxito de la red neuronal en la detección de patologías.

El estudio de MRI destacó la validación clínica exhaustiva en un entorno médico real, lo que sugiere un enfoque más completo para evaluar la eficacia del modelo. Esto puede ser especialmente importante en la adopción del modelo en la práctica médica.

La colaboración con expertos en radiología en el estudio de MRI podría haber aportado conocimientos médicos adicionales y contribuido a la mejora del modelo. La combinación de experiencia clínica y capacidades de IA es una tendencia en auge en la

¹¹ MRI, Magnetic resonance imaging, trad. *Imagen por resonancia magnética*.

¹² DSC, coeficiente de Sørensen-Dice, una medida estadística que compara la similitud de dos muestras basada en los elementos que tienen en común.

¹³ HGG, Glioma de alto grado.

¹⁴ LGG, Glioma de bajo grado.

investigación médica.

Ambos estudios podrían mencionar la necesidad de mantener el modelo actualizado en función de las cambiantes prácticas médicas y los avances tecnológicos en el campo de la imagenología.

En última instancia, el análisis comparativo destaca la similitud en los enfoques y logros entre los dos estudios, lo que sugiere que la detección de tumores cerebrales utilizando redes neuronales convolucionales es un campo en desarrollo con enfoques comunes y resultados prometedores. Sin embargo, las diferencias en aspectos como la validación clínica y la colaboración con expertos pueden influir en la confiabilidad y aplicabilidad del modelo en un entorno médico real.

Aunque el estudio ha demostrado resultados prometedores en la detección de tumores cerebrales utilizando redes neuronales convolucionales (CNN) con la arquitectura VGGNet16, también es importante tener en cuenta las limitaciones que podrían afectar la generalización y aplicabilidad del modelo en un entorno clínico:

- **Tamaño y Diversidad del Conjunto de Datos:** El tamaño del conjunto de datos utilizado puede afectar la capacidad del modelo para generalizar de manera efectiva a casos no vistos. Además, la diversidad de las imágenes en términos de género, edad, etnia y otras características demográficas puede influir en la capacidad del modelo para adaptarse a diferentes poblaciones.
- **Variabilidad de las Imágenes:** Las imágenes médicas pueden variar en calidad, iluminación, posición del paciente y artefactos. Si el modelo no se entrena con una amplia variedad de condiciones, podría tener dificultades para manejar estas variabilidades en un entorno clínico real.
- **Sobreaprendizaje (Overfitting):** Aunque el EarlyStopping se utilizó para evitar el sobreaprendizaje durante el entrenamiento, es posible que el modelo todavía haya memorizado características específicas del conjunto de datos de entrenamiento y no generalice bien a datos nuevos. Esto podría ser especialmente relevante si el conjunto de datos no es lo suficientemente grande o diverso.
- **Interpretabilidad del Modelo:** Las CNN, especialmente modelos profundos como VGGNet16, pueden ser inherentemente difíciles de interpretar. Esto puede ser un desafío en un entorno médico donde los médicos necesitan entender y justificar las decisiones tomadas por el modelo.

- **Etiquetado de Datos:** La precisión de las etiquetas de las imágenes en el conjunto de datos es crucial. Si hay errores en las etiquetas o si las imágenes se clasificaron incorrectamente, el rendimiento del modelo podría verse afectado.
- **Limitaciones Éticas y Legales:** La implementación de un modelo de diagnóstico automatizado en un entorno clínico plantea cuestiones éticas y legales. Los médicos deben ser cautelosos al confiar completamente en un modelo sin la supervisión humana adecuada.
- **Heterogeneidad de Tumores:** Los tumores cerebrales pueden variar en tipo, tamaño, forma y ubicación. El modelo podría tener dificultades para detectar tumores menos comunes o con características atípicas.
- **Sensibilidad a Cambios en los Protocolos de Imagen:** Los protocolos de adquisición de imágenes pueden variar entre instituciones médicas y con el tiempo. El modelo puede no rendir tan bien si se presenta con imágenes adquiridas con protocolos diferentes a los utilizados en el conjunto de datos de entrenamiento.
- **Necesidad de Validación Clínica:** Aunque el modelo tiene un alto rendimiento en términos de métricas de evaluación, es esencial realizar estudios clínicos y validaciones en un entorno médico real antes de implementar el modelo en la práctica clínica.
- **Actualización Continua:** Los modelos de IA, incluidas las CNN, pueden volverse obsoletos con el tiempo debido a avances tecnológicos y cambios en la adquisición de imágenes y protocolos clínicos. El modelo requerirá una actualización constante para mantener su precisión y relevancia.

Si bien las redes neuronales convolucionales y la arquitectura VGGNet16 ofrecen avances significativos en el análisis de imágenes médicas, es fundamental considerar estas limitaciones y desafíos para garantizar un despliegue efectivo y seguro en un entorno clínico.

4. Conclusiones

El estudio demuestra la eficacia de la arquitectura VGGNet16 en la clasificación de imágenes de radiografías cerebrales con y sin tumor cerebral. Las siguientes conclusiones se pueden derivar de los resultados y el proceso llevado a cabo:

- **Potencial de las CNN en Imágenes Médicas:** La utilización de redes neuronales convolucionales, en particular VGGNet16, ha demostrado su potencial para mejorar la precisión y eficiencia

en el diagnóstico radiológico de patologías cerebrales. La capacidad de aprender representaciones jerárquicas de las imágenes resulta valiosa en la detección de patrones y estructuras complejas.

- Importancia del preprocesamiento: El preprocesamiento adecuado de las imágenes es esencial para garantizar que el modelo reciba datos coherentes y relevantes. La normalización, eliminación de artefactos y redimensionamiento de las imágenes contribuyeron al rendimiento del modelo.
- Ajustes de Hiperparámetros: La elección adecuada de hiperparámetros, como el Learning Rate Schedule y EarlyStopping, es crucial para un entrenamiento eficaz y para evitar el sobreajuste del modelo.
- Alto Rendimiento del Modelo: El modelo entrenado alcanzó una alta exactitud del 98.68% en la predicción de imágenes de prueba, lo que sugiere que es capaz de identificar la presencia de un tumor cerebral en una radiografía cerebral con un alto grado de certeza.
- Validación y Aplicación Clínica: La validación del modelo utilizando el conjunto de imágenes separadas al comienzo del estudio y la comparación con la matriz del archivo CSV respaldan la capacidad del modelo para realizar predicciones precisas. Esto podría tener aplicaciones clínicas valiosas como una herramienta de apoyo para radiólogos en la detección temprana de tumores cerebrales.

El estudio demuestra que la combinación de la arquitectura VGGNet16, técnicas de preprocesamiento y ajustes cuidadosos de hiperparámetros puede resultar en un modelo de CNN altamente preciso para la detección de tumores cerebrales en imágenes de radiografías. Los resultados respaldan la viabilidad de esta metodología en la práctica médica y resaltan el papel prometedor de la inteligencia artificial en el campo de la radiología.

Referencia

- [1]
- [2]
- [3]

[4]

[5]

[6]

[7]

[8]

[9]

[10]

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei., *Desafío de reconocimiento visual a gran escala de ImageNet*. IJCV, 2015.

[12] J. Raja, Shanmugam, P. Pitchai, and R., "An automated early detection of glaucoma using support vector machine based visual geometry group 19 (vgg-19) convolutional neural network," *Wireless Personal Communications*, vol. 118, pp. 523–534, 2021.

[13] S. R. Shah, S. Qadri, H. Bibi, S. M. W. Shah, M. I. Sharif, and F. Marinello, "Comparing inception v3, vgg 16, vgg 19, cnn, and resnet 50: A case study on early detection of a rice disease," *Agronomy*, vol. 13, no. 6, p. 1633, 2023.

[14] R. Cuenca, "La génesis del uso de las radiaciones en la medicina," *Colombia Médica*, vol. 28, no. 1, pp. 34–41, 1997.

[15] M. R. Villafuerte and A. Martínez-Dávalos, "El uso de los rayos x en la medicina," *Boletín de la Sociedad Mexicana de Física*, vol. 9, no. 4, pp. 213–218, 1995.

[16] M. Alcaraz Baños, A. Gmez Moraga, M. Dato Gmez, J. Navarro, and M. Canteras Jornada, *Efecto genotxico inducido por la exposicin a rayos X durante exploraciones complejas de radiodiagnstico médico*. Oncología (Barc.), 2002.

[17] H. L. K. H. H. U. Z. T. Haq, Ejaz Ul; Jianjun, "An mri-based deep learning approach for efficient classification of brain tumors," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 6, pp. 6697–6718, 2023.