

## **Clasificación de tumores cerebrales mediante redes neuronales VGG16, InceptionV3 y ResNet50. Análisis comparativo**

Classification of brain tumors using VGG16, InceptionV3 and ResNet50 neural networks.  
Comparative analysis

Jesús Alejandro Hidalgo Gámez <sup>1\*</sup> <https://orcid.org/0009-0006-7094-0999>

Daylenn Magalys Rodríguez Pacheco <sup>1</sup> <https://orcid.org/0009-0002-0284-2829>

Arturo Orellana García<sup>1</sup> <https://orcid.org/0000-0002-3652-969X>

<sup>1</sup> Universidad de las Ciencias Informáticas, Carretera a San Antonio de los Baños, Km 2 ½, reparto Torrens, municipio Boyeros, La Habana, Cuba.

\*Autor para la correspondencia. ([jesusahg@estudiantes.uci.cu](mailto:jesusahg@estudiantes.uci.cu))

## RESUMEN

El cáncer constituye una de las primeras causas de muerte a nivel internacional. Un tumor cancerígeno, es una formación anómala de células que puede llegar a diseminarse por todo el cuerpo produciendo grandes daños a la salud, y eventualmente, la muerte. Con el objetivo de apoyar el diagnóstico y la clasificación de este padecimiento tan delicado se realizó una investigación científica, que compara el desempeño de tres de las redes neuronales convolucionales más influyentes de los últimos tiempos en el análisis de imágenes médicas. Para determinar cuál de estas es la más indicada para apoyar a la plataforma de visualización de imágenes médicas XAVIA PACS-RIS, del Centro de Informática Médica. Con el apoyo de herramientas como TensorFlow, Keras y Google-colaboratory, se utilizaron redes preentrenadas de estas tres arquitecturas para trabajar sobre el análisis del Dataset seleccionado de Kaggle, luego de seleccionar los hiperparámetros correspondientes, entrenar y probar los modelos, se determina con un 98.5% de *Accuracy*, así como otras métricas favorables, a la red neuronal VGG16 como las más óptima para esta tarea.

**Palabras clave:** red neuronal convolucional; tumor cerebral; VGG16; XAVIA PACS-RIS.

## ABSTRACT

Cancer is one of the leading causes of death internationally. A cancerous tumor is an abnormal formation of cells that can spread throughout the body, causing great damage to health and, eventually, death. With the aim of supporting the diagnosis and classification of these conditions, this scientific investigation presented, compares the performance of three of the most influential convolutional neural networks in the analysis of medical images of recent times, to determine which of these is the most suitable to support the XAVIA PACS-RIS medical image visualization platform, from the Medical Informatics Center. With the support of tools such as TensorFlow, Keras and Google-colaboratory, pre-trained networks of these three architectures were used to work on the analysis of the selected Kaggle Dataset, after selecting the corresponding hyperparameters, training and testing the models, it is determined with a 98.5% *Accuracy*, as well as other favorable metrics, show the VGG16 neural network as the most optimal for this task.

**Keywords:** brain tumor, convolutional neuronal network, VGG16, XAVIA PACS-RIS.

Recibido: 01/07/2024

Aceptado: 03/07/2024

## Introducción

El cáncer se ubica entre las primeras causas de muerte tanto a nivel internacional como en Cuba. Un tumor puede producirse inicialmente en cualquier parte del organismo, incluso en el cerebro, aunque es poco común en comparación a otras localizaciones; también cabe la posibilidad de que producto de una metástasis proliferen en esa parte del organismo. Un tumor cerebral, ya sea canceroso o no, afecta gravemente al portador indistintamente de cuál sea su posición y la función que se encuentre comprometida (Cáncer 2022).

El diagnóstico de estas patologías está estrechamente relacionado al análisis de imágenes, ya sean resonancias magnéticas o tomografías computarizadas, el uso de la tecnología es crucial. Actualmente se han desarrollado diversas soluciones informáticas que responden a necesidades médicas tanto para la gestión de centros de salud o la atención sanitaria como para apoyar a los doctores en sus diagnósticos. Estos softwares se sostienen en distintas ramas de la informática, ejemplo el programa desarrollado por la Universidad de los Andes en 2021 basado en tecnologías de aprendizaje automático, una parte de la inteligencia artificial; que es capaz de detectar tumores cerebrales, localizarlos para el usuario y clasificarlos entre tres tipos por medio de resonancias magnéticas de cuatro clases (García Ruiz, Sarmiento Barón 2021). En busca de una mayor eficiencia en los servicios de salud trabaja el Centro de Informática Médica (CESIM) de la Universidad de las Ciencias Informáticas (UCI), el cual cuenta entre sus productos con XAVIA PACS RIS, una plataforma por medio de la cual se gestionan las tareas del departamento de radiología en los hospitales en que se encuentra en uso. El programa permite la visualización de las

imágenes médicas, así como un análisis más detallado de las mismas (XAVIA PACS RIS Sistema de información radiológica. Manual de Usuario 2014; XAVIA PACSViewer Versión 3.2 Estación de diagnóstico general. Manual de usuario 2017).

La plataforma logra su objetivo de facilitar el trabajo de los profesionales de los departamentos de radiología, pero carece de algunas herramientas que pudiesen agilizar considerablemente los diagnósticos y más en casos en que el tiempo no está del lado del paciente.

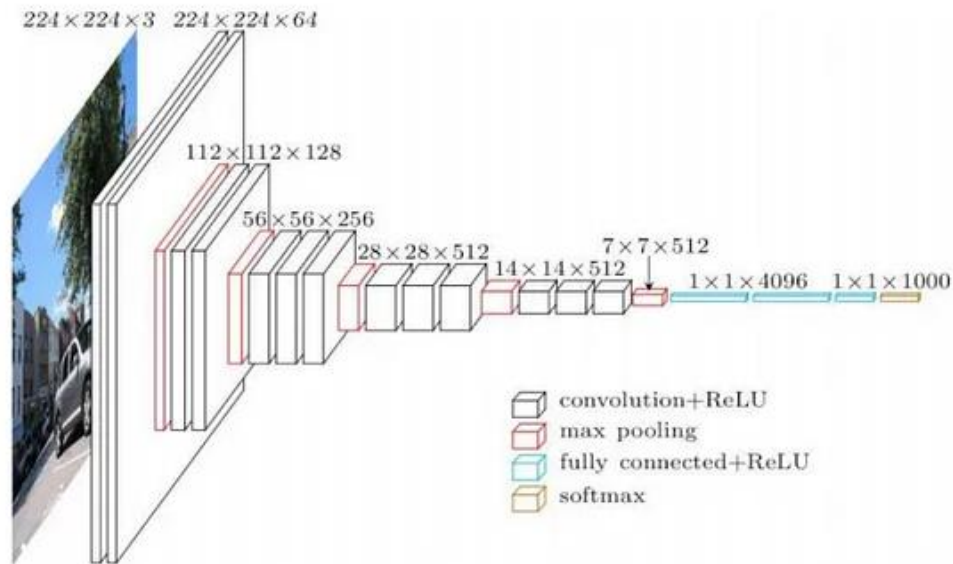
Problema de investigación: ¿Cómo contribuir al diagnóstico y el análisis de tumores cerebrales en el módulo dedicado a la visualización de imágenes médicas XAVIA PACS?

Objetivo general: Diseñar, implementar y probar una red neuronal convolucional que sea capaz de detectar y clasificar tumores cerebrales en la plataforma dedicada al análisis de imágenes médicas XAVIA PACS-RIS.

## Métodos o Metodología Computacional

### VGG16

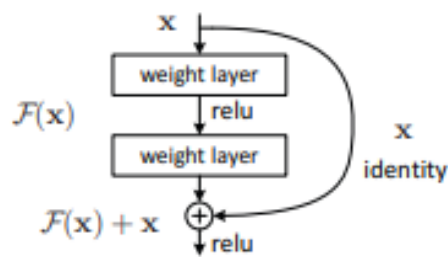
Constituye una red ampliamente conocida por la comunidad internacional, es utilizada para el reconocimiento de imágenes con regularidad debido a su puntaje de 92.7%, top cinco en *Accuracy* con el set de datos de ImageNet de 14 millones de imágenes, pertenecientes a 1000 clases en el Desafío de Reconocimiento Visual a Gran Escala de ImageNet (ILSVRC 2015). Posee una arquitectura bastante regular, con cinco bloques de capas convolucionales con pequeños filtros de 3x3, seguidas de *max-pooling* con ventanas de 2x2, termina con tres capas densas, de 4096 neuronas las dos primeras, y la última de 1000 para la clasificación, con activación *softmax* por supuesto (Thaker 2021; Daegyun 2021).



**Fig. 1-** Arquitectura VGG16. Fuente:(Thaker 2021)

## ResNet50

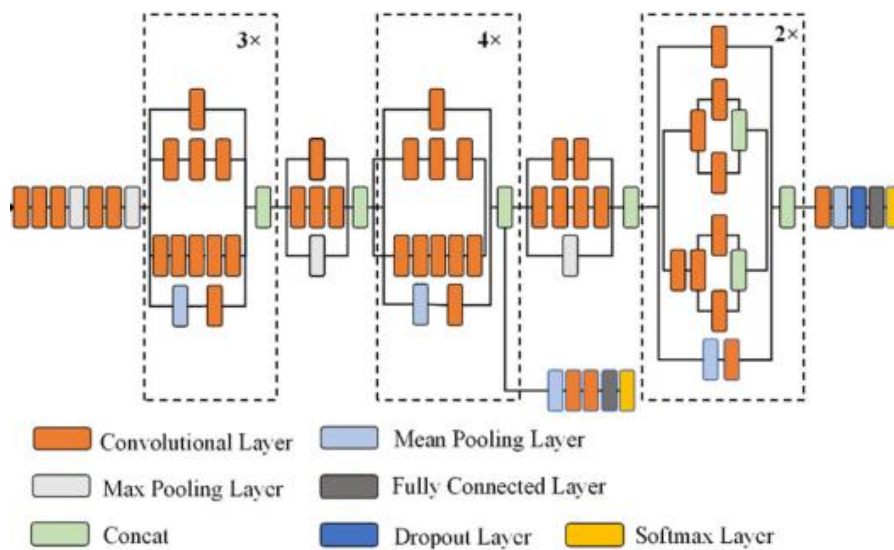
Es una red neuronal residual que alcanzó primer lugar en ILSVRC 2015, las redes de este tipo se especializan en reducir la pérdida y lograr un mejor desenvolvimiento durante el entrenamiento, ya que mantienen con mayor solidez los valores aprendidos. Es un enfoque que permite construir redes profundas sin el problema de la pérdida del gradiente, donde este se vuelve tan pequeño que los modelos dejan de aprender a fines prácticos, consiste en hacer que las salidas de neuronas en específico se salten algunas capas, de forma tal que la salida de una capa  $F(x)$ , se vuelva  $F(X) + x$ , donde la última  $x$  se refiere a las características que van a la siguiente capa pero no salieron de la anterior, propiciando que existan caminos más cortos y la carga de aprendizaje sea menor. ResNet50, debe la última cifra de su nombre a la cantidad de capas que posee, siendo 49 de estas de tipo convolución y la última, densa, para las clasificaciones (Kaiming et al. 2015; Luqman, Fady 2021).



**Fig. 2-** Bloque de aprendizaje residual. Fuente:(Thaker 2021)

### InceptionV3

La arquitectura de tipo *inception* ha evolucionado con los años logrando cada vez mejores resultados. Es un enfoque novedoso entre las CNN simples, promueve la utilización de varios bloques de convolución y *pooling* funcionando en paralelo, que vuelven a agruparse y subdividirse durante toda la red, esto permite que la red sea capaz de analizar distintas características de forma independiente, a varios niveles de profundidad, logrando análisis mucho más complejos y precisos. InceptionV3 cuenta con 20 millones de parámetros y hasta seis bloques de inception funcionando en paralelo (Luqman, Fady 2021).



**Fig. 3-** Arquitectura de InceptionV3. Fuente:(Luqman, Fady 2021)

## Dataset

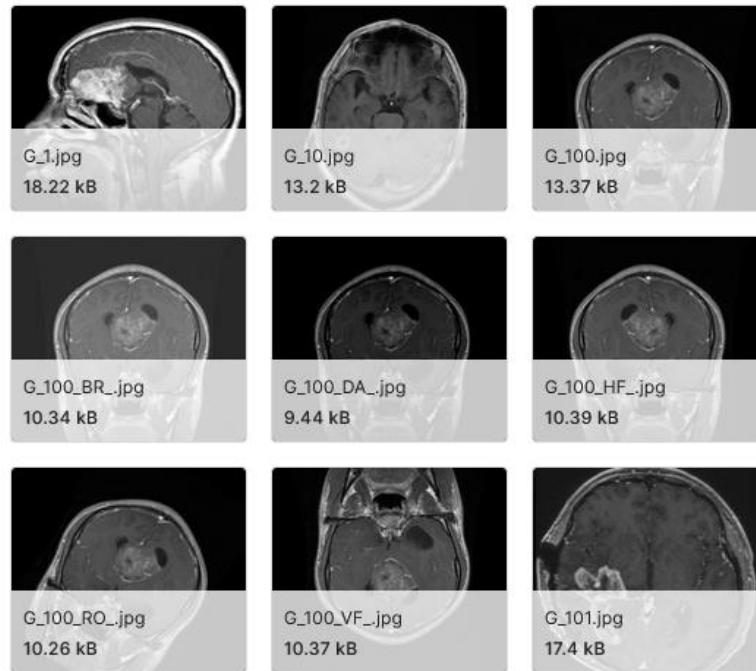
El set de imágenes utilizado durante el desarrollo de los modelos fue descargado de Kaggle, una plataforma web para la colaboración y el desarrollo de la comunidad de científicos de datos e ingenieros de aprendizaje automático a nivel mundial. Consta de 44 000 imágenes divididas en 4 categorías: "glioma", "meningioma", "pituitary" y "no-tumor", de 11 000 imágenes cada una. Como puede deducirse los datos están equilibrados, las imágenes además están limpias (sin presencia de ruido) y con las mismas dimensiones: (224,224,3). Mismas que cabe recalcar, no cumplen con los estándares internacionales para el tratamiento de imágenes médicas DICOM, sin embargo, dado que el modelo a desarrollar constituye solo una herramienta diagnóstica que ayudará a establecer una cola prioritaria de casos, utilizar un formato .JPG disminuye los tiempos de entrenamiento y los recursos de hardware consumidos durante este. Sería bastante complicado y costoso encontrar un set de imágenes tan extenso que cumpla con dichos estándares (*Tumores cerebrales - MRI - Dataset*).

Fue dividido en entrenamiento, validación y prueba para los desarrollos del modelo, de esta forma se utilizarían los primeros para entrenar el algoritmo y los datos de validación para evaluar la función de pérdida después de cada paso y actualizar los pesos con el optimizador correspondiente, al final del entrenamiento se prueba cada modelo con el último grupo de imágenes, para comprobar su eficiencia con datos con los que no entrenó.

**Tabla 1-** Distribución de los datos.

Clases	Entrenamiento	Validación	Pruebas	Total
<b>Glioma</b>	8 000	2 000	1 000	11 000
<b>Meningioma</b>	8 000	2 000	1 000	11 000
<b>Pituitary</b>	8 000	2 000	1 000	11 000
<b>No-tumor</b>	8 000	2 000	1 000	11 000

Fuente: Elaboración propia



**Fig. 4-** Muestra de las imágenes del set. Fuente: Elaboración propia.

### Google-Colaboratory, TensorFlow y Keras

Google-Colaboratory es una herramienta en línea ofrecida por Google, que permite la creación de cuadernos colaborativos, donde se puede ejecutar código de Python en la nube. Brinda muchas ventajas a investigaciones como esta, ya que permite desarrollar algoritmos potentes sin la necesidad de instalar las librerías correspondientes o tener el hardware local necesario para ello (Calvo Moratilla 2021a).

TensorFlow es una plataforma de extremo a extremo para el desarrollo de modelos de aprendizaje automático, posee varios niveles de abstracción para la adaptabilidad de los usuarios. Keras, su API de alto nivel, ofrece una experiencia introductoria a estas tecnologías, cuenta además con bibliotecas de potentes herramientas complementarias para el manejo de datos y modelos. Keras permite construir modelos de aprendizaje automático complejos con cierto grado de sencillez, y guarda modelos preentrenados de distinta índole en su módulo *applications* (Por qué TensorFlow 2024; Keras: Deep Learning for humans 2024).



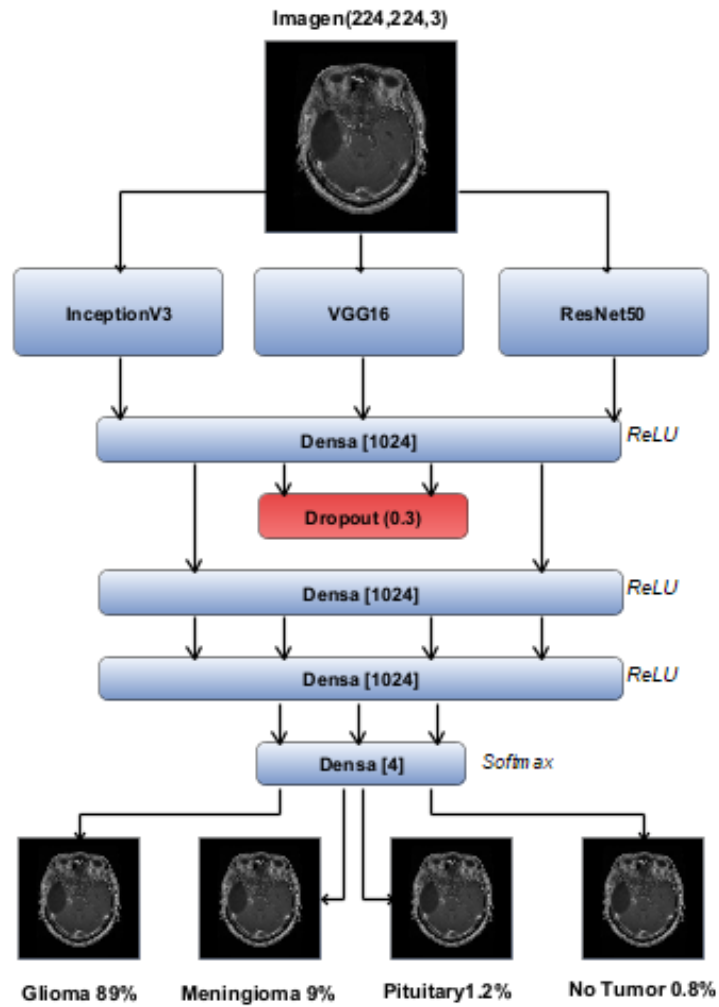
## Resultados

Se define una arquitectura inicial básica para las últimas capas de las Redes Neuronales Convolucionales seleccionadas, conformada por unas 3 capas densas de 1024 neuronas y activación ReLU (Calvo Moratilla 2021b), se aplica un *Dropout* de 0.3 después de la primera de estas (Oliva Rodríguez 2019), y por último nuevamente una capa completamente conectada con cuatro neuronas de activación *softmax* para la predicción multiclase. Para las 32 000 imágenes de entrenamiento se delimita un tamaño del *batch* de 32, durante un período de aprendizaje de 10 épocas de 1000 pasos cada una. Adam fue utilizado como optimizador para los tres modelos, ha demostrado converger con cierta rapidez, sin perder precisión, siendo actualmente uno de los mejores para la inicialización experimental (Toquero Barón 2020), y se calcula la función de pérdida con Categorical-Cross Entropy, a través de este procedimiento se calcula que tan lejos están las predicciones de cada modelo de la realidad, la seleccionada, es en específico para casos multiclase (Medina Sáez, Urra Galdames 2020). Con el fin de evaluar el modelo se utilizan principalmente Accuracy y Puntaje F1, que son funciones de alto nivel para calcular el rendimiento de modelos como este. (Salazar et al. 2018; Matamala 2022)

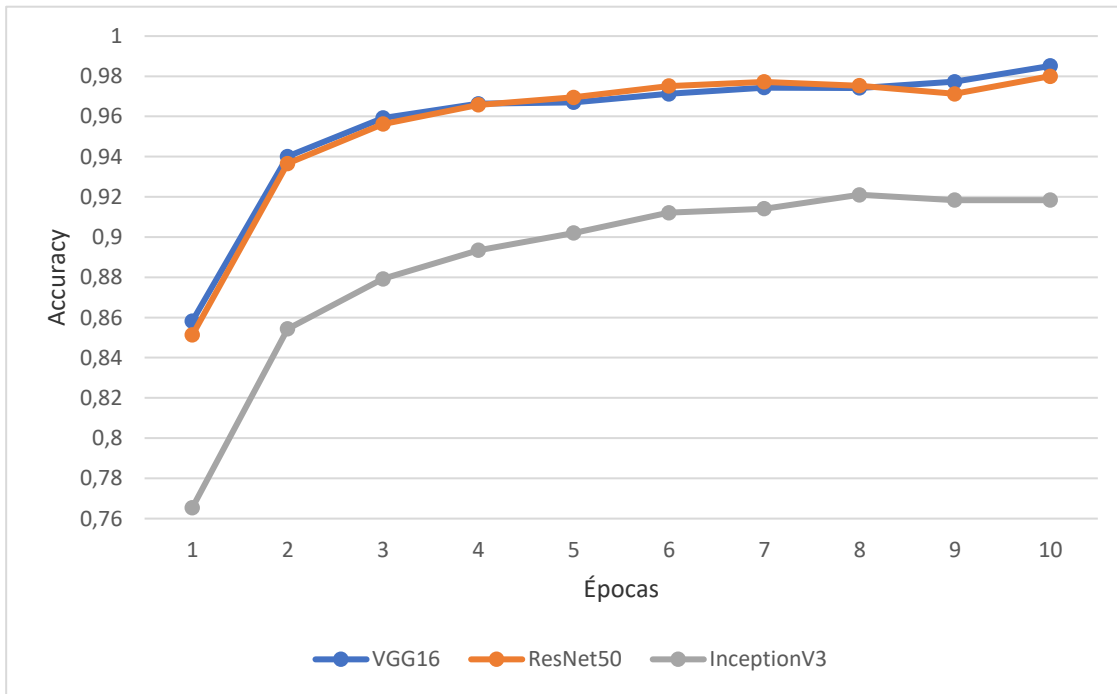
```
model.add(Dense(1024,activation='relu'))
model.add(Dropout(0.3))
model.add(Dense(1024,activation='relu'))
model.add(Dense(1024,activation='relu'))
model.add(Dense(4,activation='softmax'))

model.compile(optimizer='adam',loss='categorical_crossentropy', metrics=['accuracy',Precision(),Recall()])
```

**Fig. 5-** Fragmento de código extraído del cuaderno de Google-Colaboratory donde se definen hiperparámetros y métricas. Fuente: Elaboración Propia,



**Fig. 6-** Arquitectura de las últimas capas. Fuente: elaboración propia



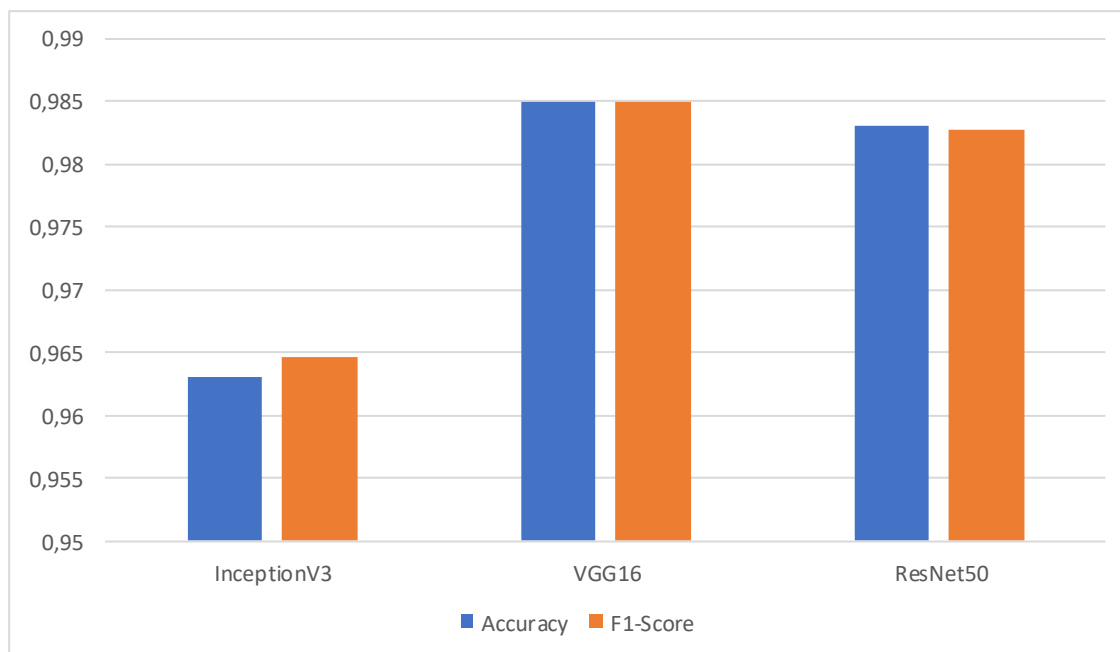
**Fig. 7-** Accuracy obtenido durante las 10 épocas de entrenamiento por los tres modelos implementados. Fuente: Elaboración propia.

**Tabla 2** Métricas obtenidas por los modelos con los 4000 casos de prueba.

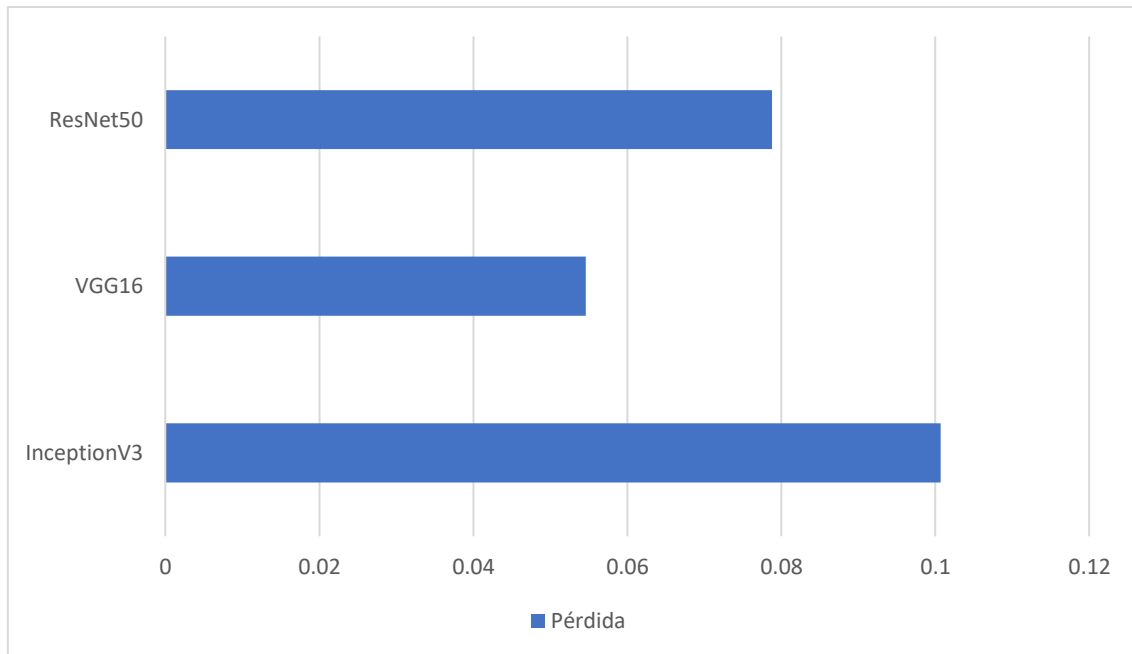
Modelo	Accuracy	Precision	Recall	Loss	F1-Score
InceptionV3	0.9630	0.9714	0.958	0.1007	0.9646
VGG16	0.9850	0.9855	0.9843	0.0546	0.9849
ResNet50	0.9830	0.9832	0.9822	0.0788	0.9827

Nota: Se resalta en rojo el modelo superior en cada métrica.

Fuente: Elaboración propia.



**Fig. 8-** Accuracy y F1-Score, obtenido por los modelos con los 4000 casos de prueba. Fuente: Elaboración propia.



**Fig. 9-** Valor de la pérdida obtenida por los modelos con los 4000 casos de prueba. Fuente: Elaboración propia.

## Discusión

Como puede observarse, el modelo VGG16 tuvo un mejor desempeño tanto durante el entrenamiento, como en el período de prueba, alcanzando un *Accuracy* del 98.5%, y una pérdida de tan solo 0.054, un resultado superior al 98.14% del modelo de clasificación binaria ResNet50 de (Deshpande, V. Estrela, Patavardhan), probablemente porque el volumen de sus datos era considerablemente menor, con tan solo 2900 imágenes. También fue superado el modelo desarrollado por (García Ruiz, Sarmiento Barón 2021), quien obtuvo un *Accuracy* del 78% con un modelo preentrenado de *InceptionV3* con 2253 imágenes. Algo distinto ocurre con (M. Sarhan 2020), donde se utilizan transformaciones de tipo Wavelet y CNN para lograr un *Accuracy* de 99.3%, dicho modelo fue entrenado con imágenes de mejor calidad: 512x512, preprocesadas por las transformaciones antes mencionadas, reduciendo la cantidad de características de menos importancia, y por consiguiente volviendo el modelo más eficaz.

Una de las principales limitantes en el desarrollo de modelos predictivos para la clasificación de imágenes es la falta de los recursos tecnológicos necesarios, recurriendo a metodologías como transfer-learning (Lubinus Badillo 2021; Carballo Pacheco 2020), que al ser modelos preentrenados, no logran la especificidad requerida en muchas ocasiones. Otras variables de importancia son la cantidad, calidad y confiabilidad de las imágenes con las que se trata, ya que, aunque existen actualmente muchas fuentes para la extracción de neuroimágenes (Yuting et al. 2022), resulta complejo encontrar bases de datos equilibradas y de buena calidad. Siendo en la mayoría de los casos la falta de imágenes, el principal obstáculo de investigaciones como esta. (Nguyen et al. 2020; Lee et al. 2020)

## Referencias

- CALVO MORATILLA, José Javier, 2021a. *Ayuda Al Diagnóstico De Cáncer De Colon Mediante Técnicas De Aprendizaje Automático*. Grado De Ingenieria Informatica. Universitat Politècnica De València.
- CALVO MORATILLA, José Javier, 2021b. *Ayuda Al Diagnóstico De Cáncer De Colon Mediante Técnicas De Aprendizaje Automático*. Grado De Ingenieria Informatica. Universitat Politècnica De València.
- CARBALLO PACHECO, Juan José, 2020. *Clasificación De Imágenes Médicas Con Técnicas De Deep Learning*. GRADO EN INGENIERÍA INFORMÁTICA EN INGENIERÍA DEL SOFTWARE. UNIVERSIDAD DE EXTREMADURA.
- DAEGYUN, Choi, 2021. UAV-Driven Structural Crack Detection And Location Determination Using Convolutional Neural Networks. *Sensors*.
- DESHPANDE, Anand, V. ESTRELA, Vania Y PATAVARDHAN, Prashant. The DCT-CNN-Resnet50 Architecture To Classify Brain Tumors With Super-Resolution, Convolutional Neural Network, And The Resnet50. *Neuroscience Informatics*. P. 2021.
- GARCÍA RUIZ, Juan Camilo Y SARMIENTO BARÓN, Sofía, 2021. *Aplicación De Las Técnicas De Aprendizaje Automático Para La Ubicación Y Detección Temprana De Tumores Cerebrales* [En Línea]. Phd Thesis. Universidad De Los Andes. Recuperado A Partir De:

<https://Repositorio.Uniandes.Edu.Co/Bitstream/Handle/1992/55455/25383.Pdf?Sequence=1&Isallowed=Y>  
[Accedido 22 Junio 2023].

KAIMING, He Et Al., 2015. Deep Residual Learning For Image Recognition. [En Línea]. Recuperado A Partir De: <https://Arxiv.Org/Pdf/1512.03385>

Keras: Deep Learning For Humans, 2024 [En Línea]. Recuperado A Partir De: <https://Keras.Io/> [Accedido 3 Junio 2024].

LEE, Changhwan Et Al., 2020. Classification Of Femur Fracture In Pelvic X-Ray Images Using Meta-Learned Deep Neural Network. *Scientific Reports*. Vol. 10, N.º 1, P. 13694. DOI 10.1038/S41598-020-70660-4.

LUBINUS BADILLO, Federico, 2021. Redes Neuronales Convolucionales: Un Modelo De Deep Learning En Imágenes Diagnósticas. Revisión De Tema. *Rev. Colomb. Radiol.*

LUQMAN, Ali Y FADY, Alnajjar, 2021. Performance Evaluation Of Deep CNN-Based Crack Detection And Localization Techniques For Concrete Structures. *Sensors*.

M. SARHAN, Ahmad, 2020. Brain Tumor Classification In Magnetic Resonance Images Using Deep Learning And Wavelet Transform. *Biomedical Science And Engineering*.

MATAMALA, Nicolás Alfonso Bettancourt, 2022. *CLASIFICACIÓN DE IMÁGENES MÉDICAS UTILIZANDO REDES NEURONALES*. Memoria De Título Presentada A La Facultad De Ingeniería De La Universidad De Concepción Para Optar Al Título Profesional De Ingeniero Civil Biomédico. Chile: Universidad De Concepción.

MEDINA SÁEZ, Ignacio Y URRA GALDAMES, - Arturo, 2020. *Análisis Comparativo E Implementación De Modelos Para El Análisis De Imágenes Médicas Enfocado A Detección De Tumores Pulmonares* [En Línea]. Para Obtener El Grado De Ingeniero Civil En Informática. Chile: Universidad Del Bío-Bío. Recuperado A Partir De:

[http://Repubib.Ubiobio.Cl/Jspui/Bitstream/123456789/3796/1/Medina\\_S%C3%A1ez\\_Ignacio.Pdf](http://Repubib.Ubiobio.Cl/Jspui/Bitstream/123456789/3796/1/Medina_S%C3%A1ez_Ignacio.Pdf)

NGUYEN, Dat Tien Et Al., 2020. Ultrasound Image-Based Diagnosis Of Malignant Thyroid Nodule Using Artificial Intelligence. *Sensors*. Vol. 20, N.º 7, P. 1822. DOI 10.3390/S20071822.

OLIVA RODRÍGUEZ, Alejandro, 2019. *Desarrollo De Una Aplicación De Reconocimiento En Imágenes Utilizando Deep Learning Con Opencv*. Grado En Ingeniería Informática. Universitat Politècnica De València.

Por Qué Tensorflow, 2024 [En Línea]. Recuperado A Partir De: <https://www.tensorflow.org/about?hl=es> [Accedido 3 Junio 2024].

SALAZAR, Juan Et Al., 2018. Automatic Segmentation Of Epidural Hematomas Using A Computational Technique Based On Intelligent Operators: A Clinical Utility. *Archivos Venezolanos De Farmacología Y Terapéutica*. Vol. 37, N.º 4, P. 12.

THAKER, Tanmay, 2021. VGG 16 Easiest Explanation. *Nerd For Tech* [En Línea]. 13 Agosto 2021. Recuperado A Partir De: <https://medium.com/nerd-for-tech/vgg-16-easiest-explanation-12453b599526> [Accedido 3 Junio 2024].

TOQUERO BARÓN, Miguel, 2020. *Clasificación De Imágenes Médicas De Rayos-X Mediante Redes Neuronales Convolucionales*. Grado En Ingeniería Informática Mención Computación. Universidad De Valladolid.

Tumores Cerebrales - MRI - Dataset, [En Línea]. Recuperado A Partir De: <https://www.kaggle.com/datasets/gonzajl/tumores-cerebrales-mri-dataset> [Accedido 3 Junio 2024].

YUTING, Xie Et Al., 2022. Convolutional Neural Network Techniques For Brain Tumor Classification (From 2015 To 2022): Review, Challenges, And Future Perspectives.

### **Conflicto de interés**

Los autores de este artículo declaran que no existen conflictos de intereses asociados a la investigación científica presentada.

### **Contribuciones de los autores**

Conceptualización: Daylenn M. Rodríguez Pacheco y Jesús A. Hidalgo Gámez, Arturo Orellana García

Curación de datos: Daylenn Magalys Rodríguez Pacheco y Jesús Alejandro Hidalgo Gámez

Análisis formal: Daylenn Magalys Rodríguez Pacheco y Jesús Alejandro Hidalgo Gámez



Adquisición de fondos: Arturo Orellana García.

Investigación: Daylenn Magalys Rodríguez Pacheco y Jesús Alejandro Hidalgo Gámez

Metodología: Jesús Alejandro Hidalgo Gámez, Arturo Orellana García

Administración del proyecto: Jesús Alejandro Hidalgo Gámez

Recursos: Arturo Orellana García.

Software: Jesús Alejandro Hidalgo Gámez

Supervisión: Arturo Orellana García.

Validación: Jesús Alejandro Hidalgo Gámez, Daylenn Magalys Rodríguez Pacheco

Visualización: Daylenn Magalys Rodríguez Pacheco

Redacción – borrador original: Daylenn Magalys Rodríguez Pacheco, Arturo Orellana García.

Redacción – revisión y edición: Jesús Alejandro Hidalgo Gámez, Arturo Orellana García.

### **Financiación**

La investigación que da origen a los resultados presentados en la presente publicación recibió fondos de la Oficina de Gestión de Fondos y Proyectos Internacionales bajo el código PN305LH013-067.