

**UNIVERSIDAD NACIONAL DE EDUCACIÓN A DISTANCIA  
UNIVERSIDAD COMPLUTENSE DE MARDID**

# **Visión por computadora**

**Proyecto**

**Detección de tumores**

**Alumno: Nelson Alfonso Beas Ham**

**Matricula: 1942687**

**Fecha de entrega: 29/01/2024**

## **Introducción**

La clasificación automática de imágenes médicas es una de las áreas más relevantes en el campo de la visión por computadora, especialmente en el diagnóstico de enfermedades complejas, como los tumores cerebrales. La detección temprana de tumores en imágenes de resonancia magnética (RM) puede salvar vidas al permitir un tratamiento más rápido y eficaz. En particular, los tumores cerebrales como gliomas, meningiomas y tumores pituitarios requieren un análisis preciso para diferenciar entre estos y los tejidos cerebrales sanos.

La clasificación de imágenes de resonancia magnética cerebrales presenta desafíos debido a la variabilidad en las apariencias de los tumores y la calidad de las imágenes, lo que hace necesario el desarrollo de modelos de aprendizaje automático robustos y precisos. En este contexto, las redes neuronales convolucionales (CNN) han demostrado ser herramientas poderosas para la tarea de clasificación de imágenes, especialmente cuando se dispone de grandes volúmenes de datos para entrenar los modelos.

## **Planteamiento del problema**

La clasificación de tumores cerebrales a partir de imágenes médicas presenta varios retos técnicos y prácticos. Uno de los mayores desafíos es la correcta detección de tumores en imágenes de resonancia magnética, dado que las características visuales de los tumores pueden variar significativamente entre pacientes, afectando la precisión de los modelos predictivos. Además, la presencia de clases desbalanceadas, como la diferencia en la cantidad de imágenes entre los tumores y el tejido no tumoral, puede dificultar el entrenamiento de modelos precisos.

El modelo de red neuronal convolucional (CNN) propuesto en este estudio tiene como objetivo abordar estos problemas mediante el aprendizaje de patrones complejos en las imágenes de RM, para detectar y clasificar correctamente los diferentes tipos de tumores. Sin embargo, la pérdida creciente durante el entrenamiento, como se ha observado en algunos experimentos previos, indica que podrían existir problemas en la configuración del modelo, el tratamiento de los datos o el diseño de la red. Por lo tanto, es esencial realizar ajustes en el preprocesamiento de datos, en la arquitectura de la red y en la optimización del modelo para mejorar su desempeño y obtener resultados confiables.

El objetivo de este trabajo es desarrollar y evaluar un modelo de red neuronal convolucional capaz de clasificar de manera efectiva las imágenes de resonancia magnética en sus respectivas categorías, a la vez que se exploran posibles soluciones a los problemas mencionados, como el sobreajuste y la pérdida de precisión. Se buscará

también la implementación de técnicas como Grad-CAM para visualizar cómo el modelo realiza sus predicciones, proporcionando mayor interpretabilidad al proceso de clasificación y ayudando a los médicos a comprender las decisiones del modelo.

## Dataset

El conjunto de datos utilizado en este estudio es un subconjunto de imágenes de resonancia magnética (RM) cerebral, que incluye varias clases de tumores cerebrales. Este dataset es esencial para entrenar modelos de aprendizaje automático destinados a la clasificación automática de imágenes médicas, específicamente en la identificación de tumores cerebrales.

El dataset proviene de un conjunto público de imágenes de resonancia magnética que contiene cuatro clases principales de imágenes. Cada clase corresponde a un tipo específico de tumor cerebral o a la ausencia de tumor. Las clases son las siguientes:

**Glioma:** Tumores cerebrales originados en las células gliales, que son las células de soporte del cerebro. Los gliomas pueden variar en agresividad y presentan características que pueden ser difíciles de diferenciar visualmente en las imágenes.

**Meningioma:** Tumores que se originan en las membranas que recubren el cerebro y la médula espinal (meninges). Estos tumores pueden tener una apariencia distintiva en las imágenes de RM, pero su identificación precisa puede ser desafiante.

**No-Tumor:** Este grupo incluye imágenes de cerebros sanos, sin la presencia de tumores. La clasificación correcta de las imágenes sin tumor es crucial para evitar falsos positivos, es decir, la identificación incorrecta de un tumor donde no existe.

**Pituitary:** Tumores que se desarrollan en la glándula pituitaria, que es responsable de la producción de hormonas en el cuerpo. Estos tumores pueden afectar a diversas funciones biológicas y requieren una identificación precisa para determinar el tratamiento adecuado.

### Distribución de las Imágenes

El dataset contiene imágenes de resonancia magnética cerebral en blanco y negro (escala de grises), lo que ayuda a reducir la complejidad computacional y facilita la extracción de características relacionadas con la estructura del cerebro y los tumores. En el caso de este proyecto, el conjunto de datos fue reducido a un subconjunto más manejable, con el objetivo de entrenar un modelo eficiente.

## Modelo:

### Modelo de Red Neuronal Convolutacional (CNN)

El modelo utilizado es una **red neuronal convolutacional (CNN)**, que se emplea ampliamente en tareas de visión por computadora debido a su capacidad para extraer características jerárquicas de las imágenes. El modelo tiene la siguiente arquitectura:

1. **Capa Conv2D:** Tres capas convolucionales con filtros de tamaños 3x3 y funciones de activación **ReLU**. Estas capas permiten extraer características como bordes, texturas y patrones de las imágenes.
2. **MaxPooling2D:** Después de cada capa convolutacional, se utiliza una capa de **max pooling** 2x2 para reducir la dimensionalidad y destacar las características más importantes, mejorando la eficiencia computacional.
3. **Capa densa (Fully Connected):** Una capa densa de 128 neuronas con activación **ReLU** para integrar la información extraída por las capas convolucionales.
4. **Capa de salida:** La capa final utiliza una activación **sigmoide** para clasificar las imágenes en dos categorías: **tumor** (glioma, meningioma) y **no tumor** (pituitary, no-tumor).

Este modelo se entrena usando la **función de pérdida binary\_crossentropy** y el **optimizador Adam**, con el objetivo de minimizar el error de clasificación.

### Mapa de Calor (Grad-CAM)

El **Mapa de Calor Grad-CAM** (Gradient-weighted Class Activation Mapping) se utiliza para visualizar las regiones de la imagen que el modelo considera más importantes para su decisión de clasificación. Grad-CAM genera un mapa de calor que destaca las áreas relevantes de la imagen al calcular los gradientes de la salida del modelo con respecto a la última capa convolutacional.

El proceso es el siguiente:

1. Se obtiene la salida de la última capa convolutacional y las predicciones del modelo.
2. Se calculan los gradientes de la pérdida con respecto a estas salidas.
3. Los gradientes se combinan con las activaciones de la capa para generar un mapa de calor, el cual se normaliza para resaltar las áreas más significativas.

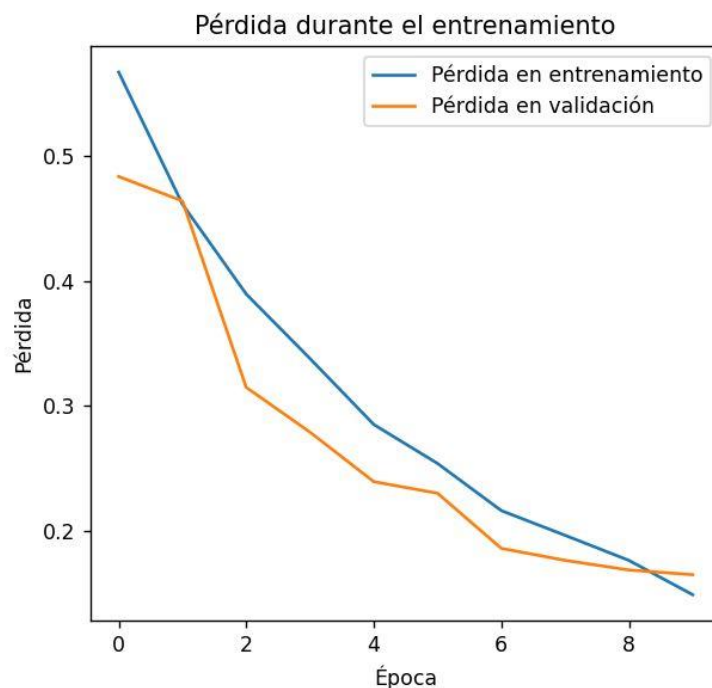
4. Finalmente, el mapa de calor se superpone sobre la imagen original para mostrar visualmente las regiones que el modelo considera cruciales para la clasificación.

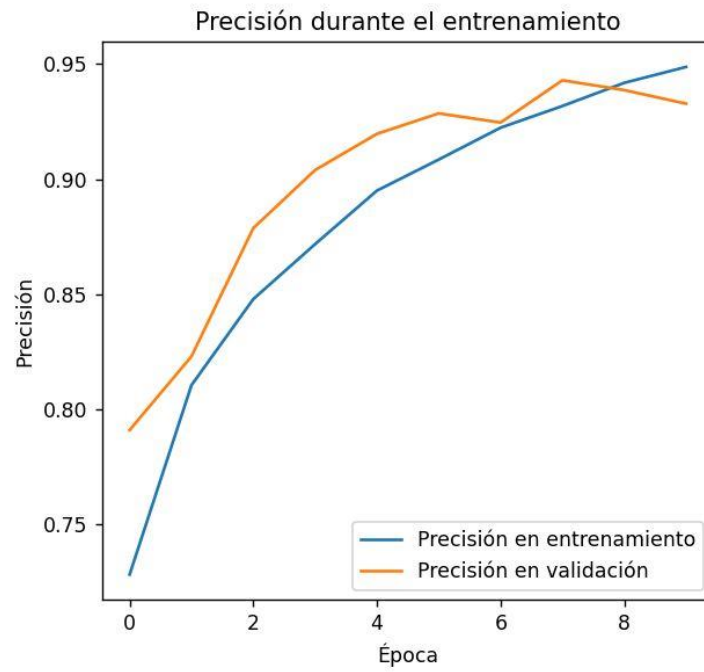
Este mapa ayuda a interpretar el comportamiento del modelo y proporciona transparencia en su toma de decisiones, lo cual es fundamental en aplicaciones médicas.

## Resultados

Tras entrenar el modelo de red neuronal convolucional (CNN) durante 10 épocas, se observó una mejora significativa en el rendimiento del modelo. La pérdida (loss) mostró una reducción constante, indicando que el modelo fue capaz de ajustarse mejor a los datos de entrenamiento y minimizar los errores en las predicciones a medida que avanzaba el entrenamiento.

Simultáneamente, la precisión (accuracy) experimentó un aumento progresivo, lo que refleja que el modelo fue cada vez más preciso al clasificar las imágenes en sus respectivas categorías de tumor/no tumor. Este comportamiento sugiere que el modelo está aprendiendo eficazmente a distinguir entre las clases, mejorando su capacidad de predicción sobre el conjunto de validación.





Ejemplos de mapas de calor generados:

Imagen Original



Tumor Detectado

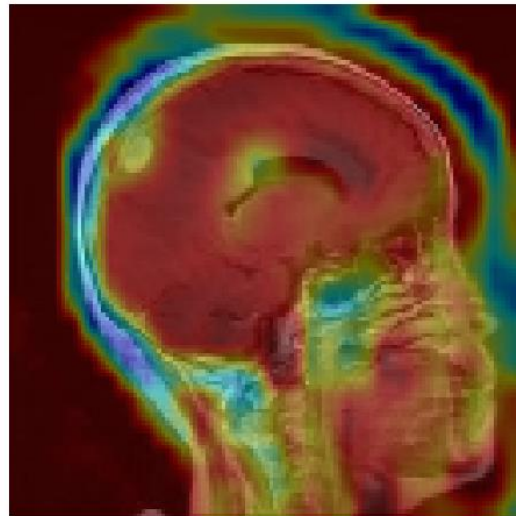


Imagen Original



Tumor Detectado

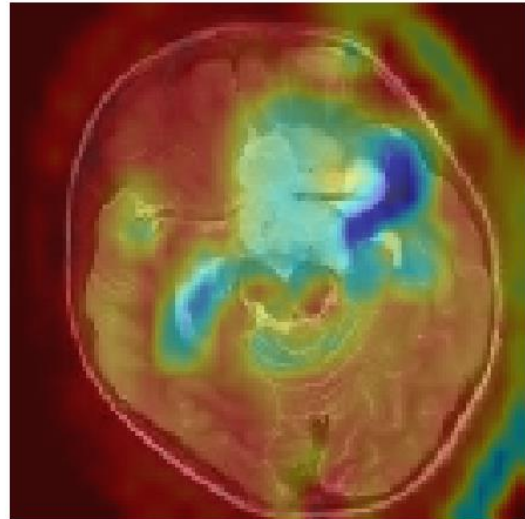
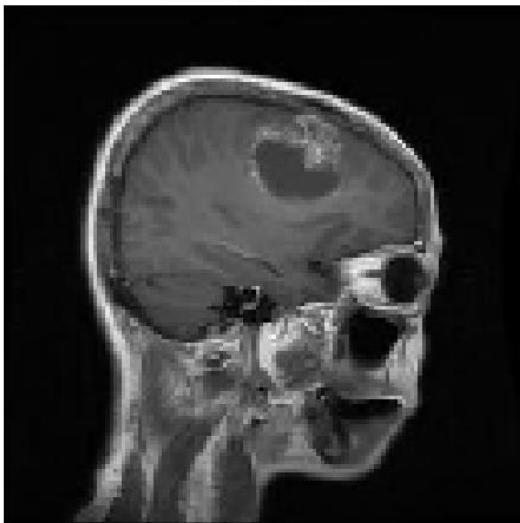
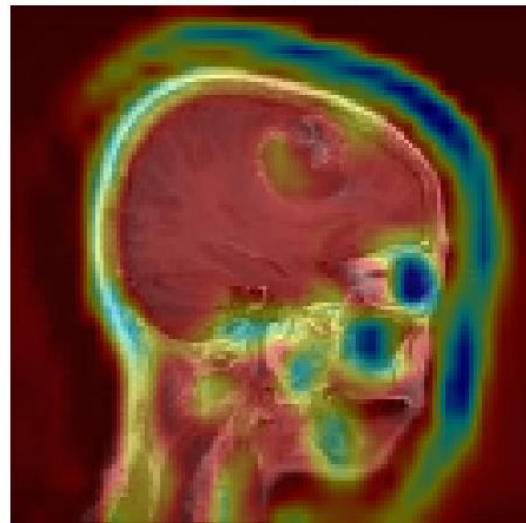


Imagen Original



Tumor Detectado



## Conclusiones

**Aplicación de técnicas de clasificación de imágenes:** A lo largo del proyecto, se ha logrado aplicar técnicas de visión por computadora y redes neuronales convolucionales (CNNs) para la clasificación de imágenes médicas, específicamente para identificar tumores cerebrales a partir de resonancias magnéticas. Se ha entrenado un modelo con distintas clases de tumores, logrando distinguir entre aquellas que corresponden a tumores (glioma y meningioma) y aquellas que no (no-tumor y pituitary).

**Importancia del preprocesamiento de imágenes:** El preprocesamiento de las imágenes, como la normalización de píxeles y la conversión a escala de grises, ha sido fundamental para optimizar el rendimiento del modelo. Esta etapa ayuda a reducir el ruido y permite

que el modelo enfoque su aprendizaje en características más relevantes de las imágenes, mejorando así la precisión del modelo. A través de la normalización de las imágenes (escala de 0 a 1), se consiguió acelerar el proceso de entrenamiento y mejorar la convergencia del modelo.

**Grad-CAM como herramienta interpretativa:** La técnica de **Grad-CAM** ha demostrado ser extremadamente útil para interpretar las decisiones del modelo, permitiendo visualizar qué áreas de la imagen son más relevantes para la clasificación. Esto no solo proporciona una mayor confianza en los resultados, sino que también facilita la comprensión de cómo el modelo realiza sus predicciones. La aplicación de Grad-CAM sobre imágenes con tumores muestra claramente las áreas específicas de la imagen que influyen en la predicción, lo cual es crucial para la validación del modelo en aplicaciones médicas.

**Desafíos y mejoras posibles:** A pesar de que el modelo mostró un rendimiento satisfactorio, se observó que la precisión puede mejorar con más datos de entrenamiento o técnicas de regularización. Asimismo, la pérdida en el entrenamiento podría ser más estable si se implementan técnicas adicionales como **data augmentation** o **early stopping**, que ayuden a prevenir el sobreajuste y mejorar la generalización del modelo.

**Importancia de la validación del modelo:** El uso de un conjunto de datos de validación ha sido clave para evitar que el modelo se sobreentrene en las imágenes de entrenamiento y, en cambio, se enfoque en aprender patrones generales que lo hagan robusto para nuevas imágenes. La mejora en la precisión y reducción de la pérdida tras 10 épocas de entrenamiento evidencian un buen desempeño del modelo en la tarea de clasificación.

## Referencias

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*, 770-778.

Selvaraju, R. R., Cogswell, M., Das, A., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 618-626.