



IA Aplicada a Ingeniería con Python
2023

Trabajo Práctico Integrador

Redes Neuronales Artificiales

Fernando Ezequiel Pose

Índice

1. Objetivos	2
2. Recomendaciones para la resolución del trabajo	2
3. Producción esperada para acreditar la actividad	2
4. Consigna	2
5. Introducción	3
6. Conjunto de datos	3
7. Clasificación de tumores en imágenes IRM	4
7.0.1. Conjuntos de entrenamiento y evaluación.	4
7.0.2. Implementación del modelo	4
7.0.3. Evaluación de resultados	5
8. Segmentación de tumores en imágenes IRM	6
8.0.1. Preprocesamiento de los datos	6
8.0.2. Implementación del modelo	8
8.0.3. Evaluación del modelo	8
8.1. Discusión y conclusiones	9
8.2. Trabajo Futuro	10
9. Material complementario	10
9.1. Base de datos	10
9.2. Notebooks	10

1. Objetivos

Ensayar la aplicación de las diferentes herramientas abordadas en cada una de las unidades del curso.

Validar los criterios asumidos para resolver problemas de la vida real con herramientas de la inteligencia artificial. Determinar la correcta vinculación entre los aspectos teóricos y prácticos.

Promover el interés respecto al estado del arte de las herramientas abordadas y sus aplicaciones prácticas actuales y/o futuras.

2. Recomendaciones para la resolución del trabajo

Evitar copiar y/o modificar soluciones de pares (compañeros, sitios de Internet, etc.), en lugar de ello, esforzarse por elaborar una producción original propia a partir del análisis y reflexión de cada una de las consignas, teniendo a mano la teoría provista en clases, la bibliografía ofrecida y todo otro material complementario que juzgue necesario para enriquecer su producción.

Reflexionar sobre los conceptos o justificaciones que se ofrecen como solución a la consigna presentada. Es decir, pueden intercambiarse opiniones, debates o puestas en común respecto a un determinado punto, pero la producción entregada debe basarse en su concepción personal del marco teórico, experiencias generales y criterios profesionales propios de su área de desempeño.

3. Producción esperada para acreditar la actividad

Presentar la documentación que considere pertinente en la tarea designada en el aula virtual de la asignatura, incluyendo un detalle de la situación problemática abordada, los objetivos estipulados y resultados logrados. Además, se deben adjuntar el material extra utilizado (datasets, códigos, etc., a fines de garantizar la repetibilidad). Realizarlo en tiempo y forma, dentro del lapso estipulado en la tarea del aula virtual.

Priorizar la calidad por sobre la cantidad, cuidando la prolijidad general en la confección, incluyendo una portada debidamente identificatoria del trabajo y las secciones propias de un informe técnico.

4. Consigna

Queda a libertad de cada cursante la propuesta de un problema (de su interés y/o necesidad) a resolver bajo las pautas del presente trabajo integrador. Dicho problema y su contexto, deberán estar aprobados por los docentes de la cátedra para garantizar, como mínimo, el alcance y complejidad pertinentes. En los casos que el presente curso sea parte de un posgrado (maestría, doctorado) se recomienda que esta aplicación sea realizada sobre una porción del trabajo de investigación de tesis.

El problema presentado, deberá modelarse como una problemática resoluble mediante herramientas de inteligencia artificial, abordadas en el curso. En caso de proponer múltiples herramientas y/o técnicas, la combinación deberá ser consistente y coherente, tanto en los aspectos teóricos como en lo propio a su implementación en software y lo que ello conlleva (eficiencia y ubicuidad, por ejemplo).

En caso de no existir una propuesta propia, puede solicitar una al plantel docente.

Si la solución abordada incluye conceptos no alcanzados en la asignatura, deberá incluir un breve marco teórico solventado por las referencias correspondientes.

Como parte de los resultados, no se pretende ninguna implementación física, ensayos o prototipo en *hardware*. La validación por simulación es más que suficiente.

5. Introducción

Un tumor cerebral es una masa o conjunto de células anormales que se desarrolla en el cerebro. Dada la rigidez del cráneo, que encierra el cerebro, cualquier crecimiento dentro de este espacio limitado puede ocasionar problemas significativos. Los tumores cerebrales pueden clasificarse en dos categorías principales: malignos (cancerosos) y benignos (no cancerosos). A medida que estos tumores crecen, aumenta la presión dentro del cráneo, lo que puede causar daño cerebral y representar una amenaza para la vida del paciente. La detección temprana y la clasificación precisa de los tumores cerebrales son esenciales para seleccionar el tratamiento más adecuado y, en última instancia, salvar vidas.

6. Conjunto de datos

El conjunto de datos utilizado en este trabajo consta de 3064 imágenes de resonancia magnética (IRM) con contraste potenciadas en T1, provenientes de 233 pacientes. Estas imágenes abarcan tres tipos de tumores cerebrales distintos:

- Meningioma (708 cortes)
- Glioma (1426 cortes)
- Pituitario (930 cortes)

El *dataset* fue proporcionado por Jun Cheng (chengjun583@qq.com), perteneciente a la *School of Biomedical Engineering Southern Medical University, Guangzhou, China*. La versión 5 del conjunto de datos utilizada puede ser descargada desde [aquí](#).

Cada imagen IRM tiene una dimensión de 512 x 512 x 4, con canales RGB y un cuarto canal de transparencia. Las imágenes están almacenadas en archivos .mat (archivos de MATLAB) y han sido preprocesadas para su uso, siguiendo el flujo de trabajo detallado en el siguiente [link](#).

A continuación, se presentan tres ejemplos de imágenes IRM junto con las máscaras correspondientes de los tumores. Estas imágenes no han sido sometidas a procesamiento adicional, excepto la conversión inicial de cada imagen contenida en el archivo .mat a una imagen y su máscara asociada.

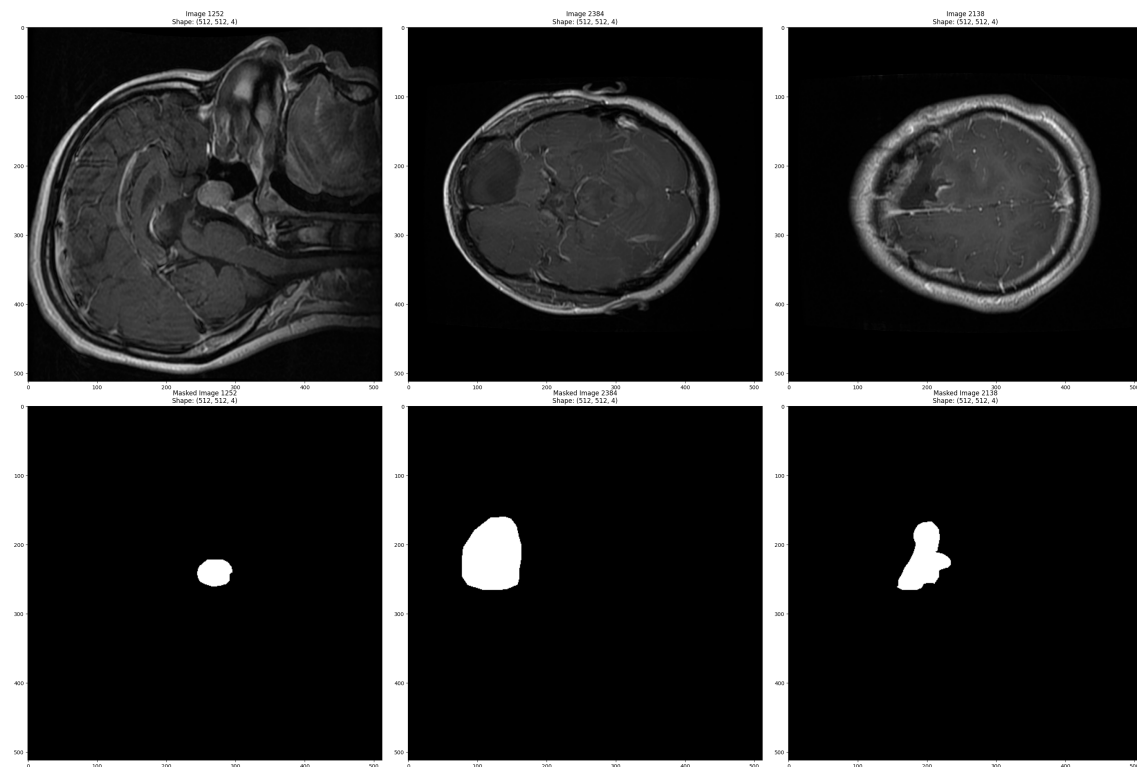


Figura 1: Ejemplo de 3 imágenes IRM con las máscaras de los tumores.

7. Clasificación de tumores en imágenes IRM

7.0.1. Conjuntos de entrenamiento y evaluación.

Como fue mencionado en la Sección 6, el conjunto de datos utilizado para el entrenamiento y la validación del modelo consta de 3064 imágenes. Estas se dividen en 708 imágenes de meningiomas, 1426 de gliomas y 930 de tumores pituitarios, cada una con su correspondiente máscara la cual indica la ubicación del tumor en el cerebro. La distribución de las muestras en los conjuntos de entrenamiento y prueba se presenta en la siguiente tabla:

Imagenes	Muestras		
	Total	Conjunto de entrenamiento	Conjunto de prueba
Meningioma	708	566	142
Glioma	1426	1140	286
Pituitario	930	744	186

Cuadro 1: Conjunto de entrenamiento y prueba

7.0.2. Implementación del modelo

Durante la implementación de la red neuronal convolucional, se adoptaron diferentes estrategias a fin de optimizar tanto el tiempo de entrenamiento como su eficiencia. A continuación, se describen las principales decisiones tomadas:

1. **Entrenamiento:** Para el entrenamiento del modelo, se realizaron 20 épocas con un tamaño de lote de 300 muestras. Además, se empleó un conjunto de validación para evaluar el modelo en cada iteración de entrenamiento.
2. **Callbacks:**
 - a) Se implementó un criterio de parada temprana con *EarlyStopping*, deteniendo el entrenamiento si la pérdida no mejoraba durante 5 épocas.
 - b) Se incorporó *ReduceLROnPlateau* para ajustar automáticamente la tasa de aprendizaje si la pérdida en el conjunto de validación no mejoraba durante 2 épocas.
 - c) Se utilizó *ModelCheckpoint* para generar puntos de control durante el entrenamiento, permitiendo la reanudación en caso de interrupciones o fallos.

La arquitectura de la red diseñada se muestra en la Tabla 2). La red incluyó cuatro capas *Conv2D* para la detección de patrones, tres capas *MaxPooling2D* para la reducción de la dimensionalidad, cuatro capas *Dropout* para la prevención del sobreajuste debido a la asignación de mayor importancia a un conjunto de neuronas sobre otras, una capa *Flatten* para redimensionar la salida a una única dimensión la cual permita ser acoplada a la última capa *Dense*, la cual se corresponde con la capa de clasificación.

Cuadro 2: Red neuronal convolucional diseñada.

Nombre (tipo) de capa	Forma de los datos de salida	Cantidad de parámetros
conv2d (Conv2D)	(None, 510, 510, 32)	896
max_pooling2d (MaxPooling2D)	(None, 255, 255, 32)	0
conv2d_1 (Conv2D)	(None, 253, 253, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 126, 126, 64)	0
dropout (Dropout)	(None, 126, 126, 64)	0
conv2d_2 (Conv2D)	(None, 124, 124, 64)	36928
max_pooling2d_2 (MaxPooling2D)	(None, 62, 62, 64)	0
dropout_1 (Dropout)	(None, 62, 62, 64)	0
conv2d_3 (Conv2D)	(None, 60, 60, 128)	73856
max_pooling2d_3 (MaxPooling2D)	(None, 30, 30, 128)	0
dropout_2 (Dropout)	(None, 30, 30, 128)	0
flatten (Flatten)	(None, 115200)	0
dropout_3 (Dropout)	(None, 115200)	0
dense (Dense)	(None, 3)	345603

La red constó de 475779 parámetros entrenables finalizando su entrenamiento en la época 15 de las 20 posibles, debido a que la pérdida en el conjunto de validación (0.1717) no tuvo una variación considerable después de la época 10.

7.0.3. Evaluación de resultados

La Figura 2 muestra las curvas de pérdida y *accuracy* durante el proceso de entrenamiento y validación en cada época del entrenamiento.

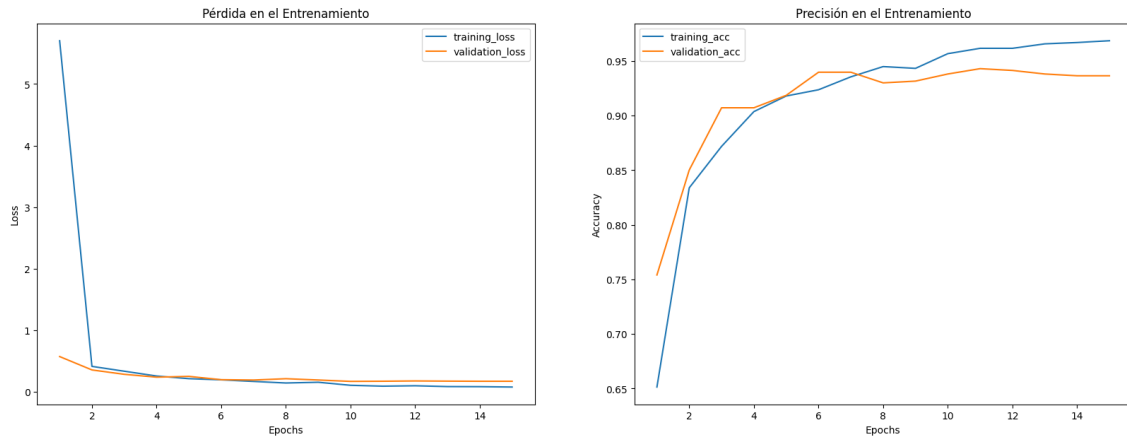


Figura 2: Curvas de pérdida y *accuracy* durante el entrenamiento.

Si bien los gráficos indican una estabilización de las curvas de *accuracy* y pérdida a partir de la época 10 y el proceso de finalización de entrenamiento terminó en la época 15, el Código 9 muestra que en la época 7 ya se había logrado un resultado acorde con valores similares de pérdida (0.19) y *accuracy* (0.93) comparables con los resultados finales. En el conjunto de prueba, se obtuvo una pérdida de 0.17 y *accuracy* de 0.93.

El tiempo total de predicción en el conjunto de evaluación fue de 6 segundos, mientras que la predicción de una imagen individual tomó 295 ms.

La Figura 3 ilustra la matriz de confusión obtenida luego de realizar la clasificación sobre el conjunto de datos de prueba.

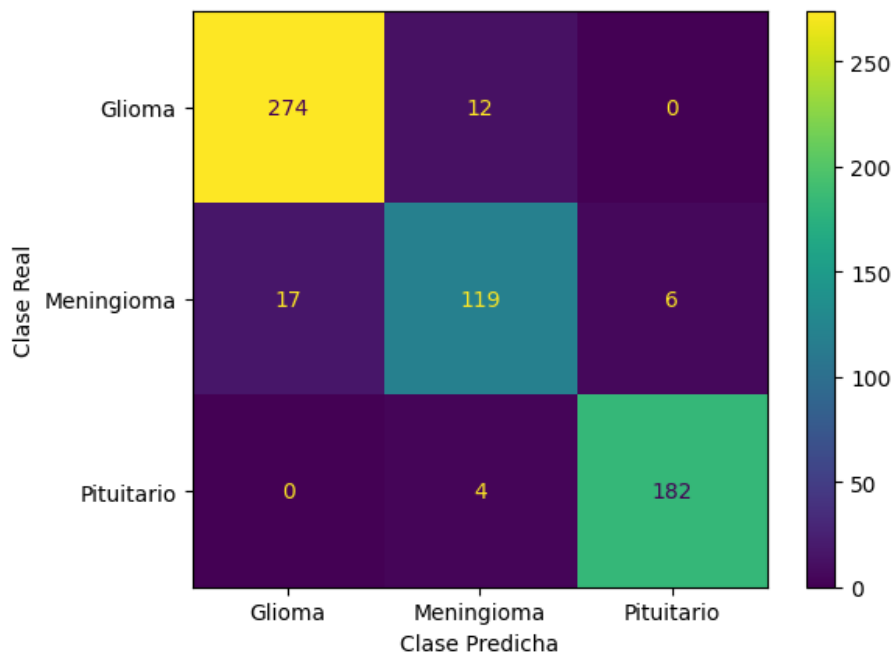


Figura 3: Matriz confusión.

A pesar de un rendimiento general sólido, la categoría Meningioma presentó mayores desafíos con 23 predicciones incorrectas, siendo principalmente confundida con la clase Glioma.

De las 614 imágenes del conjunto de datos de prueba, un total de 39 tumores resultaron mal clasificados. La Tabla 3 y la Figura 4 muestran la confianza en la predicción de 5 tumores identificados como meningiomas cuando su clase real era glioma, junto con sus respectivos identificadores de imágenes.

	Imagen	Confianza en la predicción	Clase real	Clase predicha
1	752	0.56	Glioma	Meningioma
2	756	0.65	Glioma	Meningioma
3	758	0.71	Glioma	Meningioma
4	761	0.58	Glioma	Meningioma
5	779	0.87	Glioma	Meningioma

Cuadro 3: Imágenes mal clasificadas por el modelo.

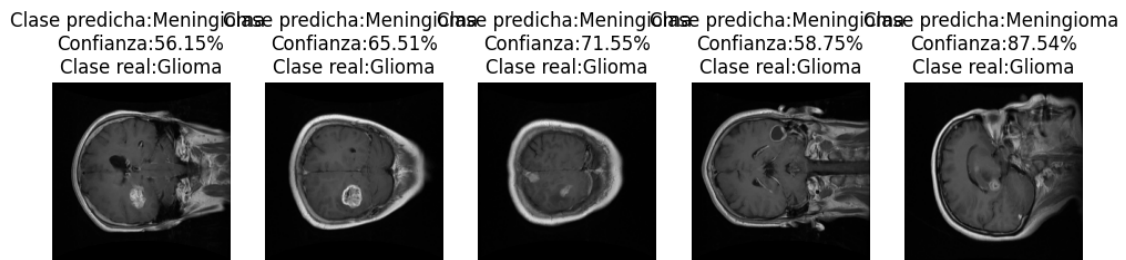


Figura 4: Imágenes mal clasificadas por el modelo.

La Tabla ?? presenta el reporte de clasificación, proporcionando detalles adicionales sobre precisión, recall y f1-score para las tres clases de tumores posibles.

	precision	recall	f1-score	support
Glioma	0.94	0.96	0.95	286
Meningioma	0.88	0.84	0.86	142
Pituitario	0.97	0.98	0.97	186
accuracy			0.94	614
macro avg	0.93	0.92	0.93	614
weighted avg	0.94	0.94	0.94	614

Cuadro 4: Red neuronal convolucional diseñada.

8. Segmentación de tumores en imágenes IRM

8.0.1. Preprocesamiento de los datos

En la etapa de preprocesamiento, se optó por una nueva resolución. Cada imagen, originalmente de dimensiones (512, 512, 4), se redujo a (128, 128, 1) píxeles, y cada máscara se ajustó a (128, 128, 1) píxeles, donde los valores de la máscara son 1 o 0. La Figura 5 muestra tres ejemplos de una imagen y su correspondiente máscara después de este proceso.

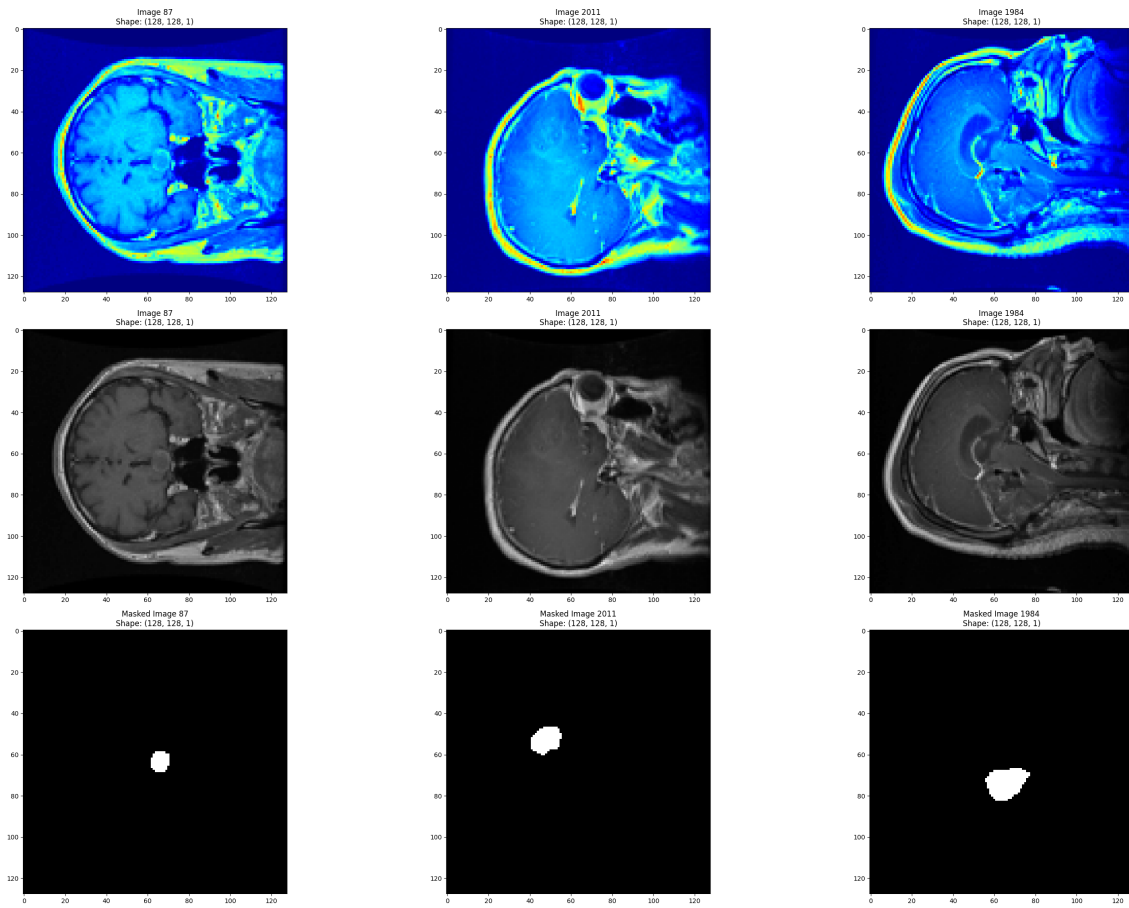


Figura 5: Ejemplos de imagen post procesamiento (128, 128, 1) y máscara (128, 128, 1)

Esta reducción de resolución contribuye a la eficiencia computacional durante el entrenamiento del modelo, disminuyendo su costo sin comprometer significativamente la calidad de la imagen.

Para la segmentación de los datos, el conjunto de datos compuesto por las 3064 imágenes y sus respectivas máscaras se dividió en un conjunto de entrenamiento, un conjunto de validación y un conjunto de prueba. La distribución de las muestras en los conjuntos se presenta en la siguiente tabla:

Cuadro 5: Conjunto de entrenamiento y prueba

	Muestras			
	Conjunto			
	Total	Entrenamiento	Evaluación	Test
Imagenes	3064	2453	488	123

8.0.2. Implementación del modelo

La arquitectura elegida para la segmentación de tumores en las imágenes es el modelo U-NET (Figura 6).

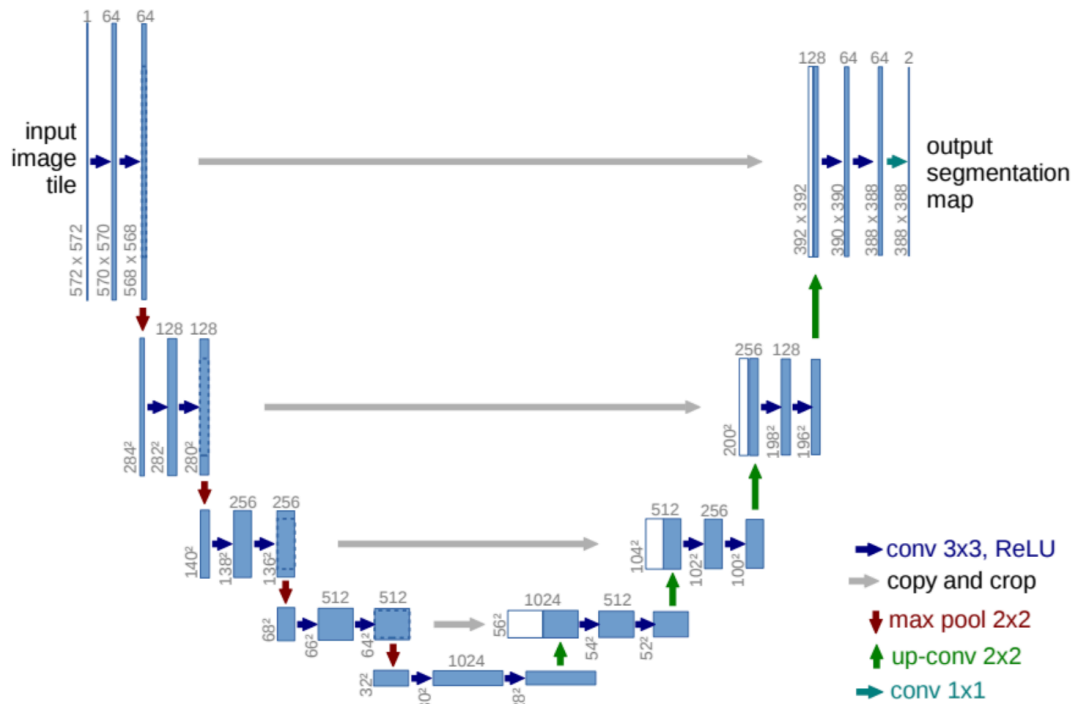


Figura 6: Arquitectura del modelo U-NET para la segmentación de tumores.

La estructura de la red se compone de una capa de entrada, 19 capas de convolución (*Conv2D*) para la detección de patrones, 4 capas de submuestreo (*MaxPooling2D*) para reducir la dimensionalidad, 2 capas *Dropout* para prevenir el sobreajuste, 3 capas de concatenación (*concatenate*) para combinar información entre capas, 3 capas de convolución inversa (*Conv2DTranspose*) para la reconstrucción de la información, y 3 capas de normalización (*BatchNormalization*) para mejorar la eficiencia del entrenamiento.

El modelo fue entrenado con un tamaño de lote (*batch_size*) de 32 y durante 40 épocas. La cantidad total de parámetros de la red fue de 8643170, de los cuales 8641186 fueron parámetros entrenables y 1984 parámetros no entrenables. El entrenamiento concluyó al finalizar la época 40. En cuanto al tiempo de entrenamiento, el modelo necesitó 605.82 segundos para entrenarse haciendo uso del *hardware* proporcionado por *Google Colab*. En el código del ejercicio (Sección 9) se puede encontrar un resumen del modelo entrenado.

8.0.3. Evaluación del modelo

En la Figura 7, se presentan las curvas de pérdida, *accuracy* y el Índice de Intersección sobre Unión (IoU) durante el proceso de entrenamiento y validación en cada época.

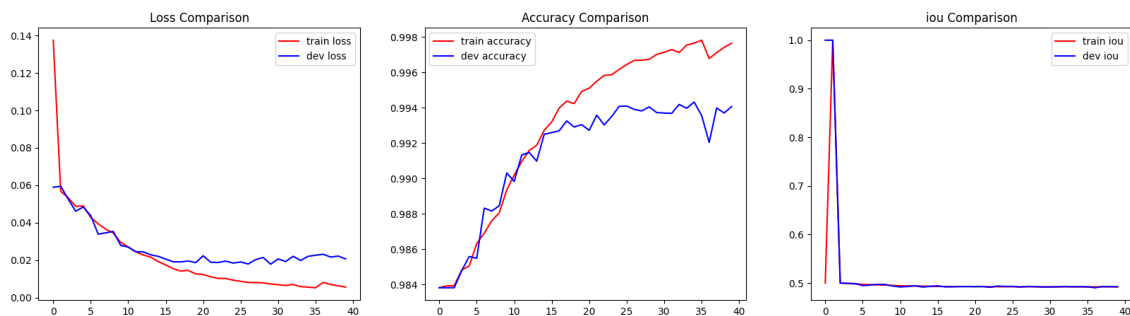


Figura 7: Curvas de pérdida, *accuracy* y IoU durante el entrenamiento.

En los gráficos, puede observarse que la métrica IoU en validación tiende rápidamente a su valor final en las primeras épocas de entrenamiento. En cuanto a las métricas de pérdida y *accuracy*, se estabilizaron alrededor de la época 35, alcanzando valores cercanos a los finales después de las 40 épocas de entrenamiento. En términos cuantitativos, se obtuvieron valores de pérdida, IoU y *accuracy* de 0.0056, 0.4921 y 0.9977 para el conjunto de entrenamiento, y 0.0206, 0.4923, 0.9941 para el conjunto de validación. Es importante destacar que la métrica IoU evalúa la precisión en la ubicación y el tamaño de la máscara predicha por el modelo en comparación con la máscara real.

Como ejemplo, la Figura 8 (izquierda) muestra cuatro imágenes las cuales fueron segmentadas por el modelo entrenado junto con sus respectivas máscaras. La misma figura, parte derecha, muestra las máscaras predichas por el modelo, la diferencia entre las máscaras reales y las predichas, y una superposición de las máscaras en las imágenes reales. Para estas cuatro predicciones, la red neuronal requirió de un tiempo medio de 0.063 segundos para predecir.

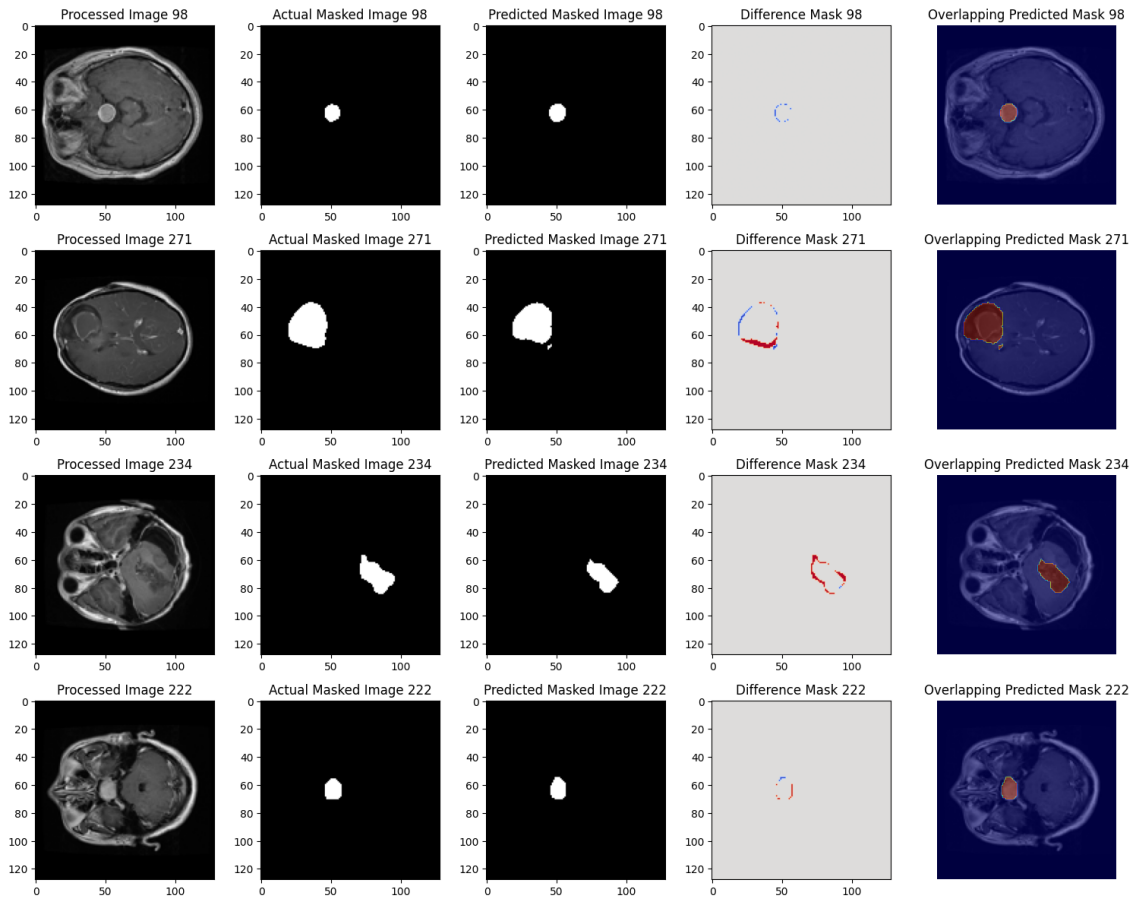


Figura 8: Resultados obtenidos en la segmentación de tumores.

Respecto al bajo valor de IoU, podría explicarse por el hecho de que, aunque la red fue capaz de detectar los diferentes tipos de tumor, la máscara predicha tiene una superficie menor que la máscara real. Por lo tanto, y mirando los resultados obtenidos (diferencia entre máscaras real y predicha), a nivel cuantitativo, para este conjunto de datos, la métrica IoU puede no ser realmente efectiva, aunque el modelo cumple el objetivo de detectar los diferentes tipos de tumor, incluso cuando el valor de IoU es bajo.

8.1. Discusión y conclusiones

La clasificación de tumores cerebrales es esencial para el diagnóstico temprano y la planificación de tratamientos adecuados. Este trabajo abordó de manera integral la clasificación y segmentación de tumores cerebrales en imágenes de resonancia magnética mediante técnicas de aprendizaje profundo. La clasificación se centró en determinar la presencia y el tipo de tumor, mientras que la segmentación se enfocó en la delineación precisa de su ubicación.

Durante el proceso de clasificación, la red neuronal convolucional entrenada demostró un rendimiento sólido, alcanzando un *accuracy* del 94 %. A pesar de estos resultados, se identificaron desafíos en la clasificación de meningiomas, que mostraron una tendencia a confusiones con la clase de gliomas. Se reconoce la necesidad de ajustes adicionales en futuras iteraciones para abordar estas cuestiones específicas.

En la etapa de segmentación, se implementó el modelo U-Net, logrando resultados prometedores. El preprocesamiento desempeñó un papel importante, y la reducción de resolución resultó ser una estrategia efectiva para mejorar la eficiencia del entrenamiento sin sacrificar significativamente la calidad de las imágenes.

Este trabajo no solo se centró al desarrollo de modelos de aprendizaje profundo, sino que también destacó la relevancia de cada etapa del proceso, desde la adquisición del conjunto de datos hasta la evaluación de los modelos entrenados. Las decisiones sobre la elección de la arquitecturas de red, las estrategias de entrenamiento y las métricas de evaluación se tomaron considerando la complejidad inherente de la tarea. En resumen, no solo se proporcionó una solución a un problema específico, sino que también se destacó la complejidad y la importancia de la intersección entre la informática y la medicina explorando formas de mejorar la práctica clínica mediante el aprovechamiento de la inteligencia artificial.

8.2. Trabajo Futuro

Como trabajo futuro, se destaca la necesidad de continuar explorando y refinando el enfoque de clasificación y segmentación propuesto con el objetivo de integrar estos modelos en entornos clínicos. A continuación, se proponen varias líneas de trabajo futuro que podrían enriquecer y expandir los resultados obtenidos:

1. Evaluación con Nuevos Conjuntos de Datos: Evaluar los modelos entrenados con nuevos conjuntos de datos con el objetivo de evaluar su capacidad de generalización y adaptación. Esto podría incluir la incorporación de imágenes sin tumores existentes.
2. Optimización de Hiperparámetros: Realizar un análisis exhaustivo de los hiperparámetros, como la cantidad de filtros, épocas y tamaño de lote, con el fin de optimizar el rendimiento de los modelos. Este proceso podría llevarse a cabo mediante técnicas de búsqueda sistemática o mediante el uso de algoritmos genéticos.
3. Interpretación de Modelos de Clasificación: Utilización de técnicas de interpretación como Grad-CAM, con el fin de visualizar y comprender las regiones de las imágenes que más influyen en las decisiones de clasificación de la red neuronal mejorando la confiabilidad del modelo.
4. Clasificación y Segmentación Integradas: Investigar la viabilidad de un enfoque unificado para la clasificación y la segmentación utilizando solo el modelo U-Net. Esto podría involucrar la utilización de cuatro clases: fondo y cada uno de los tipos de tumor posibles. La integración de estas tareas podría simplificar y optimizar los tiempos de análisis.

Estas líneas de trabajo futuro no solo ampliarían el alcance del trabajo realizado, sino que también podrían contribuir significativamente a la aplicación práctica de modelos de aprendizaje profundo en el ámbito médico, mejorando la detección y clasificación de tumores cerebrales en entornos clínicos reales.

9. Material complementario

9.1. Base de datos

- [Base de datos](#)

9.2. Notebooks

- [Armado de dataset](#)
- [Segmentación y clasificación de imágenes fmri](#)