

U-Net para la segmentación de anomalías cerebrales en imágenes médicas

VÍCTOR TRIGO

Magíster en Estadística, Universidad de Valparaíso, Valparaíso, Chile. (e-mail: victor.trigo@postgrado.uv.cl)

• **ABSTRACT** Segmentar imágenes médicas ayuda a tomar mejores decisiones médicas. Hacerlo es un trabajo arduo, tener un método de segmentación automáticos es de gran ayuda. Se propone un modelo de Deep Learning para la segmentación de anomalías cerebrales en imágenes médicas MRI. Se propone utilizar U-Net para esta tarea. Se utilizarán los datos proporcionados por un desafío en Kaggle.

• **INDEX TERMS** Deep Learning, CNN, U-Net, MRI, Kaggle

I. INTRODUCCIÓN

Una anomalía cerebral es el crecimiento anormal de algún tipo de tejido en el cerebro, que eventualmente puede ser diagnosticado como un tumor. Estos pueden clasificarse según su célula en crecimiento o según su evolución (maligna, benigna o incierta). En Chile para 2016, el diagnóstico de tumor cerebral maligno tiene una mortalidad estimada en torno a 2.6 por cada 100.000 habitantes, este número aumenta a 17.2 en pacientes de 80 años o más [1]. La detección temprana es clave para brindar un mejor tratamiento.

Para la detección de una anomalía cerebral (o tumor cerebral) se utilizan sistemas de registro de imágenes médicas. La resonancia magnética (RM) es la técnica utilizada para extraer imágenes de este tipo de tumores. Las imágenes generadas por RM realza distintos tipos de información dependiendo del tipo de filtro utilizado, los filtros son: T1, T1 Gd, T2, T2-FLAIR. Estos filtros son medidas de tiempo basado en propiedades magnéticas de la materia, ayudan a realzar tejidos blandos y duros, materia gris y materia blanca según el tiempo en la interacción con el campo magnético generado por el escáner [2].

Las imágenes médicas de RM son utilizadas por profesionales para extraer información sobre la ubicación, tamaño e impacto en el entorno de un tumor cerebral. Por otra parte, las imágenes generadas por RM contienen información de hasta 65.535 niveles de gris, lo que hace imposible para el ojo humano extraer toda la información contenida en ellas. Aquí es donde las técnicas de Deep Learning (DL) toman protagonismo, ya que sí logran capturar la información de los distintos niveles de gris, extrayendo características de forma jerárquica para luego detectar patrones, umbrales, vecindades

dentro de un espacio latente más rico en información [3].

Los modelos de DL utilizados para segmentar imágenes médicas, todos utilizan Redes Neuronales Convolucionales (CNN). A pesar de ello, los modelos son muy variados. U-Net, por ejemplo, es una estrategia de entrenamiento basado en CNN, estrategia basada en dos etapas; codificador y decodificador. El modelo obtuvo buena performance en varios problemas de segmentación de imágenes biomédicas. Destaca el rápido entrenamiento y las pocas imágenes necesitadas para su entrenamiento, 10 horas con una GPU Nvidia Titan de 6 GB.

En este proyecto se propone implementar un modelo U-Net para la segmentación de anomalías cerebrales en imágenes de resonancia magnética en filtro FLAIR.

II. MATERIALES Y MÉTODOS

II-A. DATASET

El dataset consiste en 3,929 imágenes del cerebro por resonancia magnética, junto al método de contraste FLAIR. Las imágenes serán reescaladas a 128×128 px.

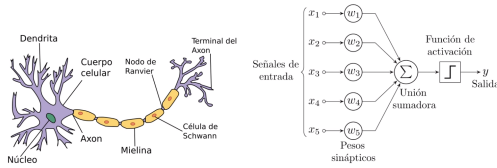
II-B. MÉTODOS

El modelo en estudio U-Net es una arquitectura de Red Neuronal Profunda, convolucional, con características de codificador y decodificador; ver [4] y [5].

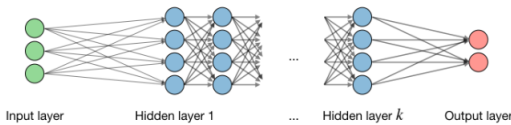
Para comprender estos modelos es necesario entender lo que es una Red Neuronal Artificial.

Inspirados en la neurona biológica, se crea un modelo probabilístico para el almacenamiento y la organización de la información en el cerebro; el perceptrón. Este da origen a la primera neurona artificial, que consta con señales de entrada

(input), pesos sinápticos (pesos), unión sumadora, función de activación y salida (output) [6].



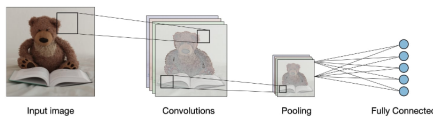
Al combinar varias neuronas artificiales, organizándolas en capas, tenemos una Red Neuronal Artificial. Está constituida por una capa de entrada (input), k capas ocultas y una capa de salida (output). Notar que en las neuronas de las capas ocultas, su input es el output de las neuronas de la capa anterior. El integrar múltiples capas ocultas se conoce como Red Neuronal Profunda, herramienta principal en Deep Learning.



Deep Learning es la disciplina dentro del Machine Learning que estudia el uso de Redes Neuronales Profundas para resolver un problema en particular: clasificación, regresión, detectar conglomerados, entre otras. Un área en que destacan los métodos de Deep Learning es en Computer Vision.

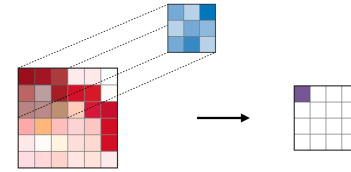
Computer Vision es dotar de visión a las máquinas, es decir, utilizar Redes Neuronales Artificiales profundas para recrear la corteza visual humana. Con esto es posible resolver problemas de clasificación, detección y segmentación de objetos en una imagen. Para esto se utilizan técnicas como: Redes Neuronales Convolucionales, Redes Neuronales Recurrentes, arquitecturas de Encoder-Decoder y Auto-Encoder y Redes Generativas Adversarias.

Redes Neuronales Convolucionales es una arquitectura de Red Neuronal Artificial Profunda, se caracteriza porque una neurona de la capa j no está conectada con todas las neuronas de la capa $j + 1$, esto disminuye la complejidad de la red (en términos de número de parámetros). Notar que el input del modelo es un arreglo rectangular de neuronas (cada neurona corresponde a un pixel de la imagen) y tres canales de color.

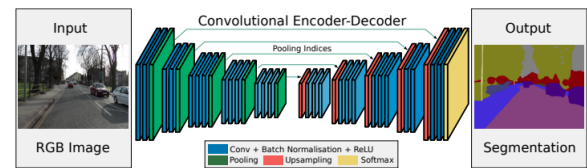


Las capas más importantes en CNN son: Convolución, Pooling y Fully Connected. Convolución es un kernel (o filtro) que recorre la imagen extrayendo características. Pooling se encarga de eliminar redundancias haciendo muestreo a través de un kernel. Finalmente, Fully Connected significa extender el resultado final en un arreglo unidimensional [7].

Los modelos en estudio tienen dos etapas: Codificador y Decodificador. Codificador, corresponde a una Red Convolu-

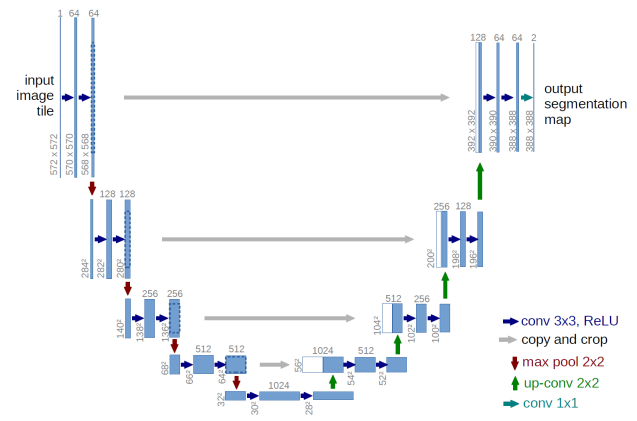


cional como ya se describió. Decodificador, corresponde a la inversa de una Convolución (convolución traspuesta). Esto se construye en forma simétrica, de manera tal que la primera convolución coincida con la última convolución traspuesta. [8], se ve así:



U-Net

U-Net es un modelo con todo lo descrito anteriormente, su nombre se debe a su forma:



El modelo consiste en dos etapas; codificador y decodificador. El codificador es la red convolucional que reduce significativamente el tamaño de una imagen y aumenta el número de canales. El decodificador es el encargado de extraer un mapa de segmentación adecuado para cada imagen. El mapa de segmentación se extrae con ayuda de una concatenación hecha desde su parte simétrica en el codificador, esta se ve en la imagen como una caja blanca. El decodificador utiliza un tipo de convolución llamada 'Convolución Traspuesta' [9].

Medida de desempeño

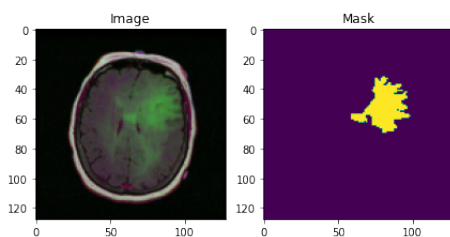
La medida de desempeño utilizada es el coeficiente Dice, se define como sigue:

$$Dice = \frac{2|A \cap B|}{|A| + |B|}$$

Donde A es la máscara de segmentación Real, B es la máscara de segmentación predicha.

III. RESULTADOS

Se seleccionaron 1,373 imágenes, cada imagen tiene una máscara de segmentación distinta de cero. Las imágenes son de tamaño 150×150 px.



El dataset se dividió de la siguiente forma; el conjunto de entrenamiento consta de 1,029 imágenes, el conjunto de validación y el conjunto de test tienen 172 imágenes.

HIPERPARÁMETROS DEL MODELO

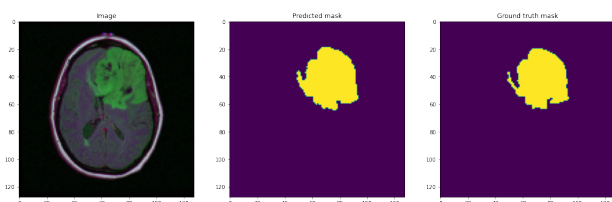
El modelo comienza con un bloque convolucional, que consta de dos convoluciones 2D con un tamaño de kernel de 3×3 y función de activación *ReLU*, cada convolución tiene un *BatchNormalization*.

El modelo U-Net propuesto toma la imagen de entrada de tamaño 150×150 (input), y las reduce a una dimensión de 16×16 , el número de filtros aumenta de 3 a 128, esto después de 4 bloques convolucionales (encoder). Luego la máscara de segmentación es extraída desde el encoder a través de 3 convoluciones traspuestas y concatenado a su simétrico en el encoder.

ENTRENAMIENTO Y DESEMPEÑO

El modelo fue entrenado en 100 épocas, con una GPU P100. Finalmente, el modelo U-Net propuesto obtuvo un coeficiente *Dice* = 0,84 como medida de generalización.

En la siguiente imagen vemos los resultados, a la derecha la imagen médica, al centro la predicción del modelo y a la izquierda la máscara de segmentación real.



IV. CONCLUSIÓN

Es posible construir un modelo U-Net para la segmentación de anomalías cerebrales en imágenes médicas. El modelo obtuvo un *Dice* = 0,84 como medida de generalización.

Se propone realizar un modelo con una función de pérdida apropiada, e insertar el coeficiente Dice en dicha función, para mejorar el desempeño.

Referencias

- [1] Contreras, L. E. (2017). Epidemiología de tumores cerebrales. Revista Médica Clínica Las Condes, 28(3), 332-338.
- [2] James, A. P., Dasarathy, B. V. (2014). Medical image fusion: A survey of the state of the art. Information fusion, 19, 4-19.
- [3] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., ... Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell, 172(5), 1122-1131.
- [4] Shen, L., Anderson, T. (2017). Multimodal brain MRI tumor segmentation via convolutional neural networks. vol. 18, 2014-2015.
- [5] Ronneberger, O., Fischer, P., Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.
- [6] Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, 65(6), 386.
- [7] Amidi, A., Amidi, S. (2018). VIP Cheatsheet: Convolutional Neural Networks. Deep Learning.
- [8] Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(12), 2481-2495.
- [9] Ronneberger, O., Fischer, P., Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

...