



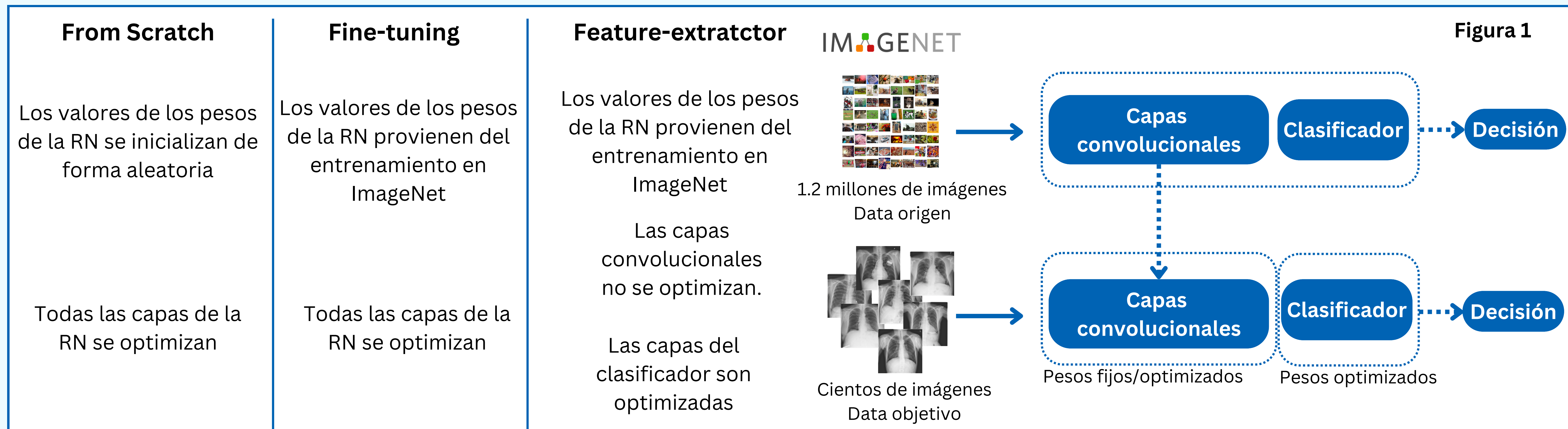
Transferencia de Aprendizaje en Redes Neuronales Convolucionales: Aplicaciones en imágenes radiológicas, limitaciones y recomendaciones

Juan Pablo Saavedra^{1 y 2}, Benjamín Pizarro³, Rodrigo Armijo^{1 y 2}, Felipe Feijoo², Hector Henríquez^{3 y 4}

1. Retina RX 2. Escuela de Ingeniería Industrial PUCV 3. Escuela de Postgrado Facultad de Medicina U. De Chile. 4. Clínica Santa María - Universidad de los Andes.

Introducción

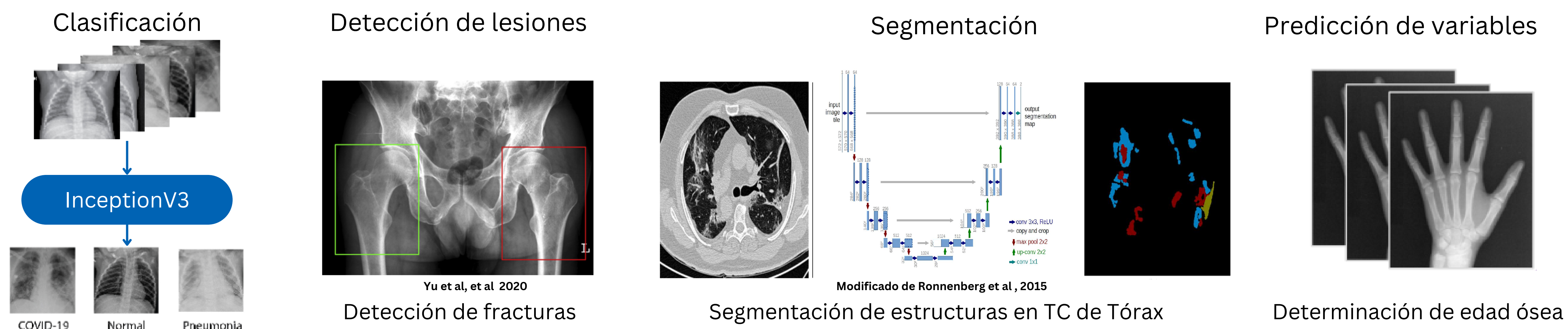
- Los modelos de Redes Neuronales (RN) y de Deep Learning (DL) han logrado equiparar y complementar la labor de radiólogos en tareas como detección o clasificación de patologías. Sin embargo para entrenar dichos modelos se requieren cientos de miles o millones de imágenes.
- La transferencia de aprendizaje (TA) es un técnica de entrenamiento de RN que permite utilizar el conocimiento que un modelo adquirió al ser entrenado en un conjunto de datos origen en una primera tarea, para ser utilizado en un conjunto de datos objetivo para una tarea diferente a la original.
- De esta forma se reutiliza la capacidad del modelo para identificar rasgos generales en las imágenes origen (bordes, patrones, texturas, o formas) y se ajusta para detectar los rasgos particulares en las imágenes objetivo (por ejemplo nódulos, fracturas, etc.)



La **Figura 1** resume de forma simplificada cómo se pueden utilizar los pesos de una RN entrenados en un conjunto de datos de imágenes naturales (datos origen), para facilitar el entrenamiento de una RN cuando se cuenta con pocas imágenes médicas para una tarea de clasificación. ImageNet corresponde una base de datos de acceso abierto que cuenta con millones de imágenes naturales etiquetadas que pueden ser usados para TA.

Aplicaciones

- Los conjuntos de datos radiológicos son escasos, difíciles de obtener y etiquetar.
- La TA permite aumentar el rendimiento de modelos de DL para tareas médicas al disminuir la necesidad de datos para el entrenamiento de RN.
- **El uso de TA puede incluso implicar reducir la necesidad de datos desde miles o millones a tan solo cientos, facilitando el desarrollo de nuevos modelos en entornos de difícil acceso a datos médicos radiológicos.**



La **Figura 2** Ejemplifica tareas médicas que pueden ser beneficiadas al utilizar transferencia de aprendizajes en modelos de DL entrenados para realizarlas

Limitaciones y recomendaciones

- Para mejorar el rendimiento de esta técnica se requiere:
 - Similitud entre datos objetivo (Imágenes médicas) y datos origen.
 - Diversidad de imágenes en el conjunto de origen.
 - Que el conjunto de datos objetivo sea pequeño.
 - Respetar que las primeras capas mantengan los pesos obtenidos tras entrenar en el conjunto de origen, el resto pueden ser inicializadas por una inicialización por distribución.
 - Aumentar el ancho o la profundidad de la red, sin embargo entrenar redes más profundas agrega complejidad y requerimiento computacional.
- Puede disminuir el rendimiento de esta técnica:
 - La existencia de cabezas auxiliares o *dropouts* en las capas de clasificación.
 - El conjunto de datos objetivo y el fuente sean muy distintos.
 - No exista suficiente variabilidad de los datos en el conjunto de origen.

Conclusiones

- La TA presenta buenos resultados en tareas médicas, como clasificación o segmentación, y se ha vuelto un estándar en el análisis de imágenes médicas.
- En un contexto en que la cantidad de datos dificulta el entrenamiento de modelos de DL, la TA logra obtener resultados comparables a los obtenidos con el entrenamiento de grandes conjuntos de datos.
- Si bien la TA tiene limitaciones, conocerlas y prevenirlas puede ayudar a mejorar el rendimiento de este tipo de técnicas usadas en modelos de DL para tareas imagenológicas y a la vez facilitar el desarrollo de nuevos modelos de DL en contexto de conjuntos de datos radiológicos reducidos.

Bibliografía

1. Morid MA, Borjali A, Del Fiol G. A scoping review of transfer learning research on medical image analysis using ImageNet. *Comput Biol Med.* 2021;128:104115. doi:10.1016/j.combiomed.2020.104115
2. Matsoukas C, Haslum JF, Sorkhei M, Söderberg M, Smith K. What Makes Transfer Learning Work For Medical Images: Feature Reuse & Other Factors. Published online October 9, 2022. Accessed July 15, 2022. <http://arxiv.org/abs/2203.01825>
3. Raghu M, Zhang C, Kleinberg J, Bengio S. Understanding Transfer Learning for Medical Imaging. Published online October 29, 2019. Accessed July 13, 2022. <http://arxiv.org/abs/1902.07208>
4. Azizpour H, Razavian AS, Sullivan J, Maki A, Carlsson S. Factors of Transferability for a Generic ConvNet Representation. *IEEE Trans Pattern Anal Mach Intell.* 2016;38(9):1790-1802. doi:10.1109/TPAMI.2015.2500224
5. Yu JS, Yu SM, Erdal BS, Demirel M, Gupta V, Bigelow M, Salvador A, Rink T, Lenobel SS, Prevedello LM, White RD. Detection and localization of hip fractures on anteroposterior radiographs with artificial intelligence: proof of concept. *Clin Radiol.* 2020 Mar;75(3):237.e1-237.e9. doi: 10.1016/j.crad.2019.10.022.
6. Ronneberger O, Fischer P, Brox T, U-Net: Convolutional Networks for Biomedical Image Segmentation. Submitted on May 18, 2015. Accessed October 18, 2022 <https://arxiv.org/abs/1505.04597>

Agradecimientos al fondo ANID Startup Ciencia: SUC210088