

Deep Learning para el diagnostico de neumonia

VÍCTOR TRIGO

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

ABSTRACT

Las enfermedades respiratorias son el tercer lugar de las causas generales de muerte en Chile, usualmente se diagnostican a través de radiografías de tórax. Mejorar nuestras habilidades de diagnostico con herramientas como maquinas de aprendizaje es lo que se propone en este articulo, en otras palabras, construir una Red Neuronal Convolutiva entrenada con imágenes de radiografías de tórax para clasificarlas los paciente entre normal o con neumonía. La arquitectura neuronal propuesta obtuvo un accuracy de 87% como medida de generalización.

INDEX TERMS

Machine Learning, Deep Learning, Artificial Neural Networks, Convolutional Neural Networks, Diagnostico Medico, Neumonia

I. INTRODUCCIÓN

Las enfermedades respiratorias son el tercer lugar de las causas generales de muerte en Chile (según el Ministerio de Salud). La neumonía es una infección del tejido pulmonar, comúnmente encontrada en niños y adultos mayores a 65 años. El diagnostico de neumonía requiere principalmente de una radiografía de tórax y de exámenes físicos. Mejorar nuestras habilidades de diagnostico es valioso y de gran valor para la medicina.

La Inteligencia Artificial (IA) tiene el potencial de revolucionar el diagnostico de enfermedades difíciles para expertos humanos [1]. Complementar el razonamiento humano con el de una maquina es necesario de cara a la innovación. Saber utilizar estas herramientas y entender sus resultados será la norma en este siglo. En este proyecto se pretende estudiar el uso de la IA (visión por computadora principalmente) en el diagnostico medico.

La visión por computadora ha incrementado su popularidad al incluir métodos de deep learning en su aplicación, principalmente con las llamadas Convolutional Neural Network (CNN). CNN es una red neuronal profunda (múltiples capas) que utilizando operaciones matriciales llamadas convoluciones es capaz de capturar información de imágenes ahorrando mucho computo respecto a las redes neuronales convencionales. Estas tienen aplicaciones muy interesantes, por ejemplo, el reconocimiento de caracteres escritos a mano [2]. Cambien en el reconocimiento de caras [3], detectar emociones y frecuencia de parpadeos (para prevenir riesgos en carretera) [4]. Los coches autónomos es lo mas reconocible

en este tema [7].

En sus aplicaciones en el diagnostico medico tenemos varios ejemplos. Para el diagnostico de cáncer de piel en la clasificación de melanoma benigno o maligno [5]. Para el diagnostico de COVID-19, se comparó radiografías de personas con COVID-19, neumonía normal y sanas, contrastando treinta y dos arquitecturas diferentes mediante aprendizaje por transferencia [6]. En el análisis de encefalograma, para detectar y clasificar epilepsia en normales, prenatales y convulsivas. [8]. Además de la detección de neumonía en radiografías [1].

En este proyecto se propone estudiar las redes neuronales convolucionales y proponer una arquitectura para modelar imágenes de radiografías de tórax y clasificarlas en si poseen o no la enfermedad de neumonía. Con el fin de conocer y aplicar la arquitectura CNN y sus componentes.

El resto del proyecto se organiza de la siguiente manera: una revisión sobre el dataset y los modelos CNN se presenta en la sección 2. Los resultados y disección de los mismos se encuentran en la sección 3. Finalmente, conclusiones se presentan en la sección 4.

II. MATERIALES Y METODOS

En esta sección se describirá el dataset utilizado y la metodología de analisis.

A. MATERIALES

El conjunto de datos es obtenido de la pagina Kaggle. En dataset consta de 5.863 imágenes de radios X de tórax

clasificados como "pneumonia" o "normal".

B. MÉTODOS

En esta sección se mostraran las técnicas utilizadas para transformar, analizar y clasificar las imágenes.

1) Preparación del Dataset

Las imágenes serán rescaledadas a una dimensión de $150 \times 150px$ y un canal de color (escala de gris), además de la normalización de los pixeles (cada pixel entre 0 y 1). También se utilizó métodos de data augmentation para balancear las imágenes entre las categorías.

2) Convolutional Neural Network (CNN)

Las Redes Neuronales Convolucionales (CNN) es una red neuronal especializada en el reconocimiento de imágenes. Intenta simular el comportamiento de la corteza visual de nuestro cerebro. Utiliza operaciones matriciales llamadas convoluciones que disminuyen la complejidad del algoritmo.

Una convolucion o Kernel, también llamada filtro, es una matriz (tamaño a elección) que se desliza por toda la imagen de entrada (matiz de pixeles) haciendo una operación matemática llamada convolucion.

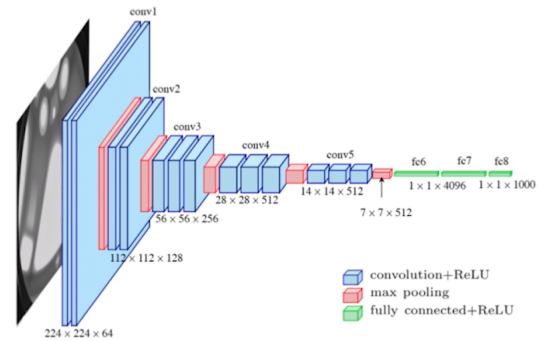
$$\begin{bmatrix} 2 & 0 & 0 & 4 & 0 \\ 1 & 1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 1 & 2 & 3 & 0 \\ 1 & 1 & 2 & 3 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 6 & 8 & 8 \\ 4 & 6 & 7 & 8 \\ 4 & 4 & 8 & 4 \\ 3 & 5 & 4 & 3 \end{bmatrix}$$

Estos se utilizan para extraer features de las imágenes, primeramente líneas y formas básicas que luego forman objetos mas complejos.

La capa de convolución posee dos hiperparametros importantes: Padding y Strides. Padding se refiere a incluir un margen de 0's en la imagen, para con esto tener mas información de los bordes de la imagen. Strides es un número que indica la cantidad de pasos que se salta la función de convolución

Otra capa importante es la llamada pooling. Es un filtro o kernel que aplica una operación matemática, en general el Max o Average. Este ayuda a disminuir el tamaño de las matrices resultantes de una convolución eliminando ruido.

Estos componentes son suficientes para crear una CNN funcional, tomando en cuenta una capa llamada Flatten que transforma el output de las convoluciones en un vector y una función de activación (como ReLu o softmax).



Existen diferentes métodos para optimizar una CNN.

1) Data Augmentation

Especialmente utilizado cuando se posee pocas imágenes o para datasets desbalanceados. Esto consiste que de una única imagen se hacen ciertas modificaciones: rotar, zoom, cambiar brillo, etc.



2) Batch Normalization

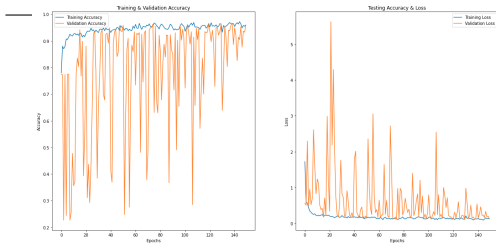
Se utiliza para reducir la escala de los pesos en una red, esto trae como consecuencia una mejor optimización respecto al descenso del gradiente. Es una capa que se aplica entre convoluciones.

III. RESULTADOS Y DISCUSIÓN

La arquitectura neuronal propuesta es la siguiente: 3 capas convolucionales con activación ReLu; 3 capas de Batch Normalization; 2 capas de MaxPooling; 2 capas de Dropout; Flatten y capa Dense con activación Softmax. Con un total de 81.890 parámetros. En resumen:

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 75, 75, 32)	320
activation (Activation)	(None, 75, 75, 32)	0
batch_normalization (Batch Normalization)	(None, 75, 75, 32)	128
conv2d_1 (Conv2D)	(None, 75, 75, 64)	18496
activation_1 (Activation)	(None, 75, 75, 64)	0
batch_normalization_1 (Batch Normalization)	(None, 75, 75, 64)	256
max_pooling2d (MaxPooling2D)	(None, 25, 25, 64)	0
dropout (Dropout)	(None, 25, 25, 64)	0
conv2d_2 (Conv2D)	(None, 25, 25, 96)	55392
activation_2 (Activation)	(None, 25, 25, 96)	0
batch_normalization_2 (Batch Normalization)	(None, 25, 25, 96)	384
max_pooling2d_1 (MaxPooling2D)	(None, 6, 6, 96)	0
dropout_1 (Dropout)	(None, 6, 6, 96)	0
flatten (Flatten)	(None, 3456)	0
dense (Dense)	(None, 2)	6914
Total params: 81,890		
Trainable params: 81,506		
Non-trainable params: 384		

El modelo fue entrenado por 150 épocas y con 40 pasos por época. El modelo obtuvo una performance de: 87.3% de accuracy y 0.4089 de perdida. Gráficamente el entrenamiento se ve de la siguiente forma:



Donde se observa un entrenamiento bastante caótico, donde tanto el set de entrenamiento como el de validación parecieran converger a un mismo resultado, con mucha mas variabilidad en el set de validación.

Al analizar la matriz de confusión tenemos:

True label	predicted label	
	Normal	Pneumonia
Normal	374	16
Pneumonia	63	171

Donde podemos obtener las siguientes métricas:

- accuracy: 0.873
- recall: 0.959
- especificidad: 0.731
- precisión: 0.856
- valor predicho negativo: 0.914

La especificidad es el valor mas bajo obtenido, lo que indica los muchos errores que comete el algoritmo al intentar clasificar la neumonía como tal.

IV. CONCLUSIONES

Construimos una arquitectura neuronal capaz de predecir la presencia de neumonía en radiografías de tórax. El modelo obtuvo un accuracy de 87% como medida de generalización. El modelo es menos fuerte en especificidad con 73% en contraste con recall de 95%.

El estudio de CNN es complejo, solamente utilizamos las características fundamentales de esta herramienta. Para trabajos futuros se propone no solo mejorar la arquitectura neuronal, sino también utilizar arquitecturas mas complejas de CNN (Alexnet, ResNet, LeNet, etc) y el uso de aprendizaje por transferencia.

REFERENCES

- [1] 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.
- [2] Durán Suárez, J. (2017). Redes neuronales convolucionales en R: Reconocimiento de caracteres escritos a mano.
- [3] Montoya, Ó. P. (2018). Redes Neuronales Convolucionales Profundas para el reconocimiento de emociones en imágenes. Universidad Politécnica de Madrid.
- [4] Pérez Rodríguez, M. (2021). Reconocimiento eficiente de caras mediante Deep Learning a partir de imágenes en el espectro visible.
- [5] Tejada Layme, G. M., & Gonzales Chama, R. P. (2020). Arquitectura de red neuronal convolucional para diagnóstico de cáncer de piel.

- [6] Lopez-Betancur, D., Bosco Duran, R., Guerrero-Mendez, C., Zambrano Rodríguez, R., & Saucedo Anaya, T. (2021). Comparación de arquitecturas de redes neuronales convolucionales para el diagnóstico de COVID-19. *Computación y Sistemas*, 25(3).
- [7] Osada Mochizuki, J. A. (2012). Implementación de un vehículo de pruebas para el desarrollo de un sistema de estacionamiento autónomo mediante visión robótica.
- [8] Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., Adeli, H. (2018). Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Computers in biology and medicine*, 100, 270-278.

...