

MACHINE LEARNING BASED PRIORITY READ LIST FOR THE DETECTION OF PNEUMONIA IN CHEST X-RAY IMAGES

Juan Irving Vasquez-Gomez^a, Hind Taud^b

^aConsejo Nacional de Ciencia y Tecnología (CONACYT)-Instituto Politécnico Nacional (IPN), Ciudad de México, México,

jvasquezg@ipn.mx

^b Instituto Politécnico Nacional (IPN), Ciudad de México, México

htaud@ipn.mx

RESUMEN

Para combatir la neumonía de forma efectiva es necesario contar con diagnósticos rápidos y precisos. Sin embargo, el número de radiólogos es limitado y suelen gastar su tiempo en revisar casos negativos. Por lo tanto, hacer una selección previa de los casos con mayor probabilidad de ser positivos para que sean analizados por los radiólogos puede agilizar el tratamiento de los casos críticos. Por otro lado, los algoritmos de visión basados en aprendizaje profundo han probado su efectividad en tareas de clasificación de imágenes. Sin embargo, a pesar de sus prometedores resultados, en el área de aplicaciones médicas, los resultados distan de ser perfectos. Por lo tanto, en este trabajo en lugar de realizar la clasificación de las imágenes utilizamos los métodos de aprendizaje profundo para establecer una lista de prioridad de lectura basada en los ‘logits’ de salida predichos por una red neuronal reciente. El método propuesto fue ajustado de acuerdo a las métricas de sensibilidad, especificidad y calificación F1 alcanzada por la red VGG en la clasificación de neumonía en el conjunto de datos ‘NIH Chest X-ray Dataset’.

Palabras claves: Neumonía, clasificación, x-ray, VGG, prioridad

ABSTRACT

To fight pneumonia fast and accurate diagnoses are required. However, the number of radiologists is limited, and their work time is usually spent in negative diagnostics. Therefore, selecting the most probable cases for being first analyzed by the radiologist can speed up the treatment of critical cases. On the other hand, deep learning algorithms have proved to be very efficient in the classification of images, however, their performance is still limited in medical applications given that they are usually pre-trained in natural environments, e.g. ImageNet, and the amount of available medical images is still small with respect to current natural environments databases. Therefore, instead of directly classifying images we use deep learning methods to provide a priority read list based on the logits inferred by a state-of-the-art neural network. Our method is set according to an evaluation of the sensibility, specificity

and F1 scores reached by VGG for the classification of Pneumonia in the NIH Chest X-ray Dataset.

Key words: Pneumonía, classification, x-ray, VGG, priority

1. INTRODUCCIÓN

Una de las acciones que mayor impacto tiene en el tratamiento de muchas enfermedades, en especial la neumonía, es la detección temprana y precisa del padecimiento. Actualmente, la detección se realiza en la mayoría de los casos apoyada por la valoración de un radiólogo que analiza una placa de rayos-x de tórax del paciente.

La problemática reside en los siguientes puntos. Primero, el número de radiólogos es pequeño comparado con la cantidad de placas que se tienen que analizar, sobre todo en regiones marginadas. Esto provoca un retraso generalizado en el proceso, generando una cola de imágenes que se tienen que analizar. Segundo, en la mayoría de los casos se trata de diagnósticos negativos, como se puede observar en la distribución de hallazgos en [1]. En consecuencia, los radiólogos gastan gran parte de su tiempo en revisar placas de casos negativos, mientras los casos positivos son postergados. Aun así, se debe recurrir al diagnóstico del experto dado que actualmente éste provee los diagnósticos más acetados y además puede proveer una explicación de su dictamen [6].

Por otro lado, las técnicas actuales de visión computacional basada en aprendizaje profundo con redes neuronales [2] han mostrado resultados importantes en tareas de clasificación, detección y segmentación de diversas entidades en imágenes. Sin embargo, su efectividad en las aplicaciones de cuidado de la salud es limitado. Dichas limitaciones se deben en parte a que el número de imágenes públicas para investigación y desarrollo tecnológico en estos temas es pequeño. Por lo tanto, se ha recurrido al uso del aprendizaje por transferencia, sin embargo, no ha sido de total efectividad debido a que las imágenes con las cuales se pre-entrenaron los modelos, son muy diferentes a las imágenes de rayos-x. e.g. ImageNet la base de datos pública más popular contiene imágenes de ambientes naturales, rurales o urbanos, los cuales distan de semejarse a las imágenes de rayos-x. En consecuencia, aplicar aprendizaje profundo directamente para resolver el problema de clasificación automática de neumonía no es recomendable, pero se pueden desarrollar métodos que asistan el diagnóstico.

En este trabajo proponemos un método para asignar a cada una de las placas radiográficas un peso para que sean ordenadas en la lista de lectura. De tal forma que los radiólogos revisen primero los casos estimados como prioritarios. Nuestro método se basa en el sistema 'Neumonitor', el cual fue desarrollado en nuestro laboratorio a partir del uso del aprendizaje por transferencia. En específico, utilizamos como base la red neuronal VGG16[2] pre-entrenada en el conjunto de datos ImageNet y cambiamos el perceptrón multicapa (MLP) de clasificación por una de las mismas

dimensiones, pero sin entrenar. En un segundo paso, entrenamos el nuevo MLP con ejemplos de la base de datos *NIH chest X-ray*[1]. Nuestro sistema fue ajustado de acuerdo con la calificación F1 y muestra prometedores resultados.

2. MARCO TEÓRICO

2.1 Aprendizaje por transferencia

El aprendizaje por transferencia reusa modelos pre-entrenados en una tarea base para resolver una tarea objetivo, donde ambas tareas comparten similitudes. En este trabajo utilizamos un modelo de red neuronal profunda para estimar la “probabilidad” de que un paciente presente neumonía a partir de la observación de su placa de rayos-x de torso. Existen al menos dos tipos de aprendizaje por transferencia en el aprendizaje profundo.

2.2 VGG16.

La arquitectura VGGNet presentada en [5], es una arquitectura de red neuronal convolucional clásica. Su particularidad se fundamenta en el concepto del uso, en dos capas sucesivas, un filtro de tamaño pequeño (3 x 3) en cada capa, en el lugar de usar, en una sola capa, un solo filtro de mayor tamaño (5x 5). Lo que implica un aumento de la profundidad de dicha red y una disminución del número de parámetros aumentando su rendimiento.

Se han desarrollados varias variantes de esta arquitectura, aumentando sus capas a 16 y 19 capas. De esta manera, el VGG16 es formada de manera sucesiva de bloques de convolución, cada bloque tiene la forma de $A \times \text{convT-P}$, donde A es el número de capas usadas, T tamaño del filtro y P la profundidad del filtro. El VGG16 es propuesto de la siguiente manera: $2 \times \text{conv3-64}$, $2 \times \text{conv3-128}$, $3 \times \text{conv3-256}$, $3 \times \text{conv3-512}$ y $3 \times \text{conv3-512}$. Finalizando con 3 capas densamente conectadas.

3. NEUMONITOR

El sistema propuesto lleva por nombre Neumonitor. Este sistema está desarrollado en tres etapas: transferencia del conocimiento, validación y predicción. Observe la figura 1. La etapa de transferencia consiste en implementar la red neuronal VGG16 e instanciarla con los pesos provistos por el paquete “Keras” opción “ImageNet”[4]. Estos pesos corresponden al entrenamiento para la clasificación de imágenes del concurso Imagenet. En la segunda etapa, ‘validación’, reemplaza el clasificador por MLP con un nuevo MLP instanciado con pesos aleatorios y se procede a entrenar la red neuronal. En la última etapa, predicción, el sistema se pone a trabajar en línea de producción, éste recibe una imagen de tipo DICOM, se extrae el mapa de bits correspondiente a la radiografía del tórax y se hace pasar por la red neuronal. La red da como resultado un valor denominado ‘logit’ para cada una de las posibles clases; dicho valor entre más se acerque a 1 es más probable que la imagen pertenezca a esa clase, es decir:

$$L(x) = \text{Sigmoid}(MLP_{\theta}(Conv_{\theta}(x))) \quad (1)$$

donde x es la imagen de entrada $Conv_{\theta}()$ es la etapa de extracción de características convolucional con los parámetros θ , $MLP_{\theta}()$ es el perceptrón multicapa con los parámetros θ y Sigmoid es la función sigmoide.

Una vez que se obtiene el valor del logit, este se asigna al archivo DICOM como un campo extra denominado 'prioridad' y se actualiza el orden de lectura.

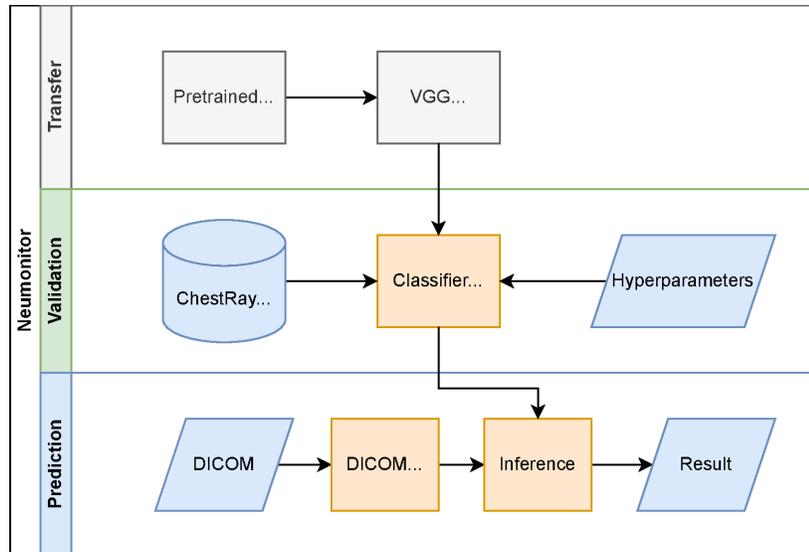


Figura 1. Sistema Neumonitor

4. EXPERIMENTOS

La red neuronal VGG16 fue implementada usando Python y el paquete de aprendizaje profundo 'Keras'. La implementación se puede descargar desde: <https://github.com/irvingvasquez/neumonitor>. Los datos de entrenamiento fueron preprocesados con los siguientes pasos: seleccionar únicamente la clase 'neumonía' y tomar ejemplos de las otras clases como una nueva clase 'no neumonía'. Balancear los datos a través de submuestreo incluyendo estratificación. Incluir aumentación de datos con espejo horizontal aleatorio, desplazamiento vertical con máximo 10% de la imagen, desplazamiento horizontal de máximo 10%, rotación aleatoria de máximo 20 grados, distorsión afina de máximo 10% y zoom de 10%.

VGG fue entrenada usando el optimizador Adam, pérdida de entropía cruzada binaria y tamaño de lote de 200 durante un máximo de 100 épocas usando terminación temprana.

Los resultados del entrenamiento se muestran en las figuras 2 y 3. Como podemos observar en las figuras, el criterio de terminación temprana detiene el entrenamiento alrededor de las 50 épocas. La terminación temprana es un método simple de regularización que cuando detecta sobreajuste de termina el entrenamiento. La

exactitud alcanzada en el conjunto de validación fue de 0.7. Se usó un umbral de 0.5 para separar predicciones positivas y negativas. Aunque pareciera que dicha exactitud es baja, cabe señalar que el objetivo no es directamente clasificar, si no, proveer un peso que estime el orden de la lista de prioridad. Es decir, en el peor de los casos los falsos negativos serán revisados después de los positivos predichos.

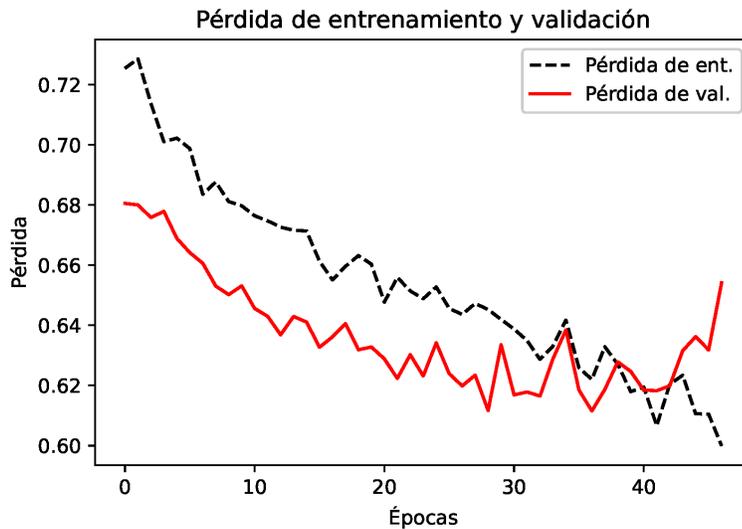


Figura 2. Gráfica del valor de pérdida durante el entrenamiento

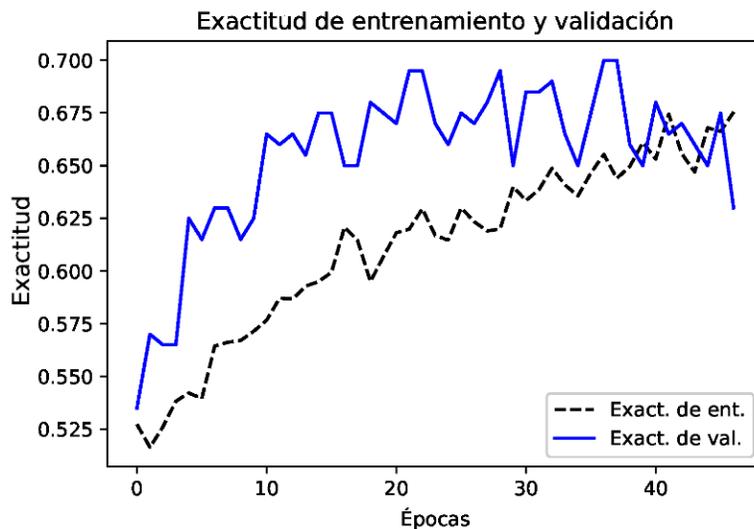


Figura 3 Gráfica del valor de exactitud durante el entrenamiento

5. CONCLUSIONES

Se presentó un método basado en aprendizaje profundo por transferencia para construir una lista de prioridad en la lectura de imágenes de rayos x en la clasificación de neumonía. El método usa los 'logits' obtenidos por la red neuronal para determinar la prioridad en la lista. El siguiente paso es realizar un estudio para garantizar que todas las imágenes sean leídas incluso aunque su prioridad sea baja.

6. REFERENCIAS

- [1] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. *IEEE CVPR* 2017.
- [2] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *Nature* 521.7553 (2015): 436-444.
- [3] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- [4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. *IEEE Computer Vision and Pattern Recognition (CVPR), 2009*.
- [5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in 3rd International Conference on Learning Re-presentations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, ConferenceTrack Proceedings (Y. Bengio and Y. LeCun, eds.), 2015.
- [6] Wu, J. T., Wong, K. C., Gur, Y., Ansari, N., Karargyris, A., Sharma, A., & Syeda-Mahmood, T. (2020). Comparison of chest radiograph interpretations by artificial intelligence algorithm vs radiology residents. *JAMA network open*, 3(10).