

# Deep Learning en Radiografía de Tórax: Detección Automática de Neumonía Mediante Redes Neuronales Convolucionales

Mariano E. Casciaro<sup>1</sup>, Hugo José  
Paladini<sup>2</sup>, Ricardo Lineros Franco<sup>2</sup>,  
Adriana Martínez<sup>2</sup>, Mario Embon<sup>2</sup>,  
Alejandro Beresñak<sup>3</sup>

1. *Instituto de Medicina Traslacional, Trasplante y Bioingeniería (IMeTTYB), Universidad Favaloro-CONICET, Buenos Aires, Argentina.*
2. *Hospital Universitario Dr. René G. Favaloro, Buenos Aires, Argentina*
3. *Investigaciones Médicas, Buenos Aires Argentina*



CONICET

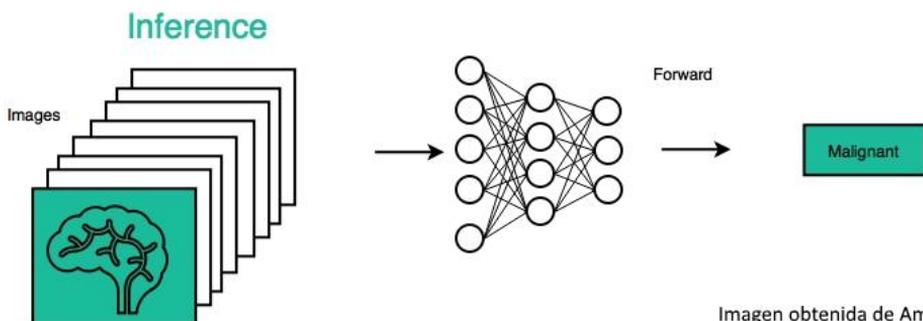
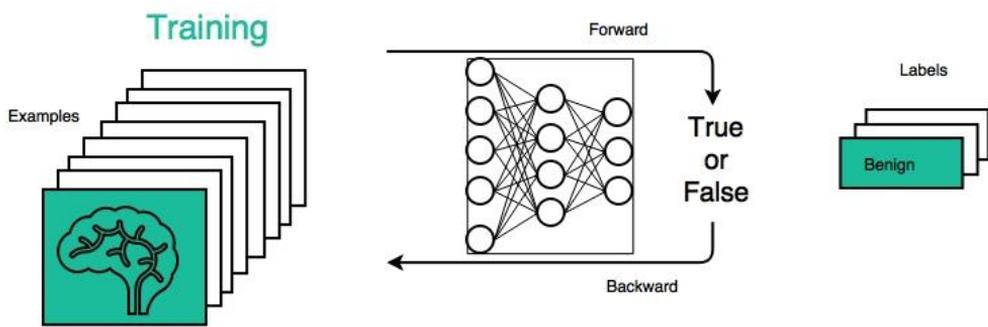


INVESTIGACIONES MEDICAS  
Centros de Diagnóstico

I M E T T Y B

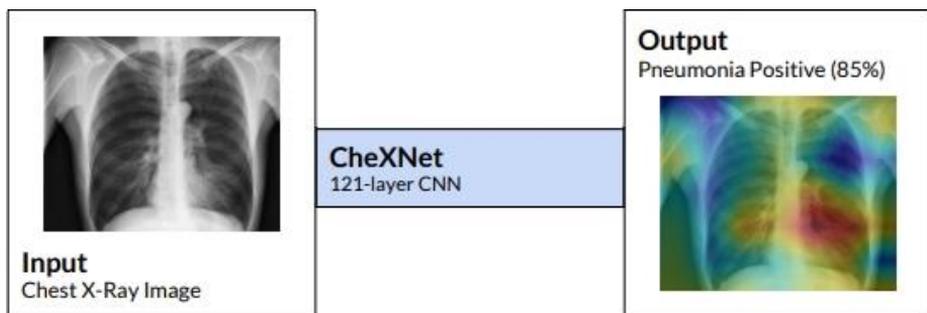
# Introducción

- Las redes neuronales convolucionales (CNN):
  - Están inspiradas en las neuronas de la corteza visual.
  - Se perfilan como el algoritmo de asistencia diagnóstica más prometedor en radiología.
  - Función de clasificación e inferencia, entrenándose con casos conocidos para predecir presencia/ausencia de enfermedades en imágenes médicas

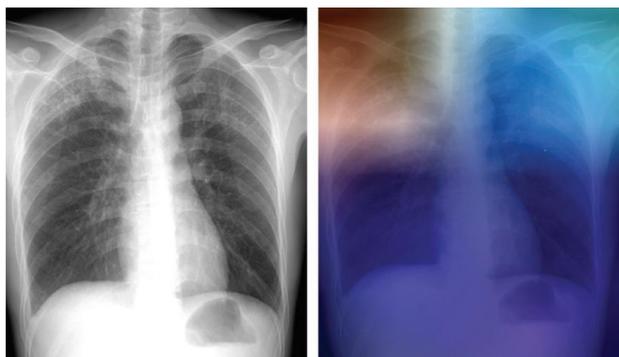


# Aplicación de las CNNs en placas de tórax

- Dos ejemplos de aplicación reciente de esta técnica:
  - Rajpurkar et al (2017):
    - Datos: **100.000 placas de tórax rotuladas** en 8 patologías no excluyentes, entre ellas neumonía.
    - CheXNet: arquitectura de red de 121 capas.
    - **Precisión del 85%**

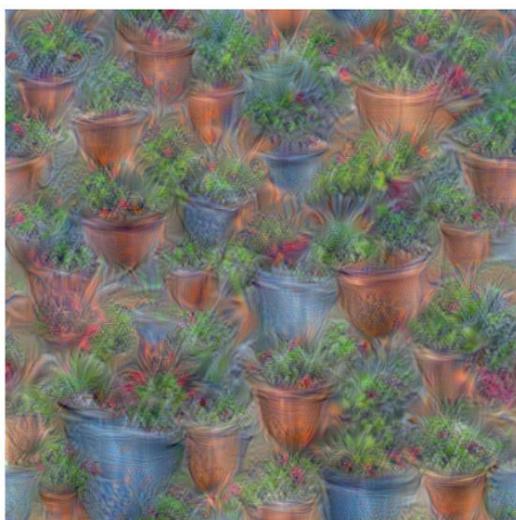


- Lakhani et al (2017):
  - Datos: 1007 placas de tórax con/sin tuberculosis.
  - AlexNet y GoogleNet: re-entrenamiento de dos redes utilizadas para clasificar objetos (no médicos).
  - **Precisión del 98%**



## Objetivos del trabajo

- Re-entrenar una red neuronal convolucional (AlexNet) para la detección de neumonía en placas de tórax.
- Determinar lo aprendido por la red en forma visual mediante la técnica de “Deep-Dreaming”.



Técnica de Deep-Dreaming para generar imágenes mediante el estímulo de las neuronas que han aprendido a clasificar (izq) gallinas y (der) macetas (imagen obtenida de <https://www.mathworks.com/help/nnet/examples/deep-dream-images-using-alexnet.html>)

# Materiales y métodos: *Dataset* del NIH



Home » News & Events » News Releases

## NEWS RELEASES

Media Advisory Wednesday, September 27, 2017

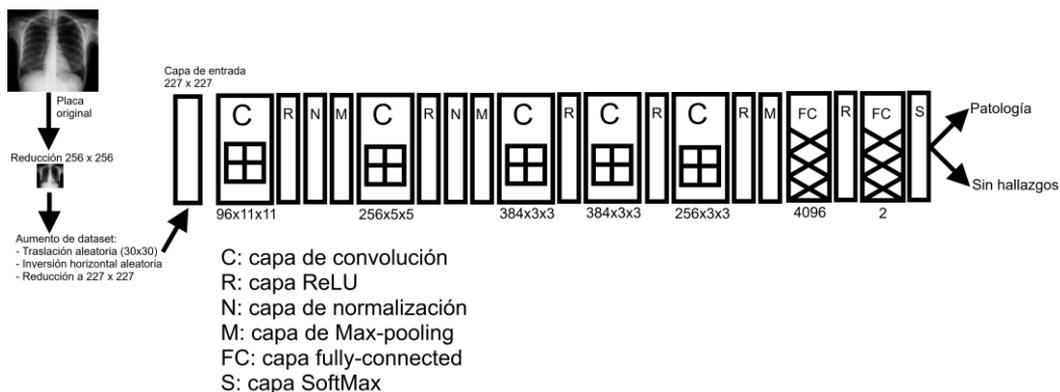
NIH Clinical Center provides one of the largest publicly available chest x-ray datasets to scientific community

*The dataset of scans is from more than 30,000 patients, including many with advanced lung disease.*

- 100.000 placas de tórax rotuladas mediante técnicas de Natural Language Processing (NLP) de los informes radiológicos.
- Debido al costo computacional, para este trabajo se descargó la mitad de la base de datos (N = 50.000).
- División en 3 conjuntos:
  - P: placas con evidencia de neumonía (N = 593)
  - NP: placas sin neumonía (N = 49.407)
  - NF: placas sin ningún hallazgo patológico evidente (N = 24.041).

# Materiales y métodos: Red CNN AlexNet

- Se utilizó un esquema de *transfer learning* modificando y reentrenando las 3 últimas capas de la red pre-entrenada AlexNet:



Arquitectura AlexNet, indicando la sucesión de capas. A la entrada puede verse el esquema de reducción de tamaño (256 x 256 pixeles) y variación aleatoria del dataset utilizado en este trabajo:

- Traslación
- Inversión horizontal
- Reducción final a 227 x 227

# Materiales y métodos: División del *dataset* y compensación del desbalance de categorías

- El desbalance en la cantidad de placas de cada categoría atenta contra el aprendizaje de la red neuronal.
- Para compensarlo, se realizó el submuestreo aleatorio de las categorías hasta alcanzar un  $N = 800$  placas para las categorías NF y NP.
- Se entrenaron dos redes, 10 veces cada una:
  - Una red para diferenciar P de NP.
  - Una red para diferenciar P de NF.
- Para cada uno de los 10 experimentos, se realizó un nuevo submuestreo de las categorías NP y NF.
- El dataset se dividió siempre en un porcentaje 80/20:
  - 80% para el grupo de entrenamiento.
  - 20% para el grupo de validación.

# Materiales y métodos: Entrenamiento de la red

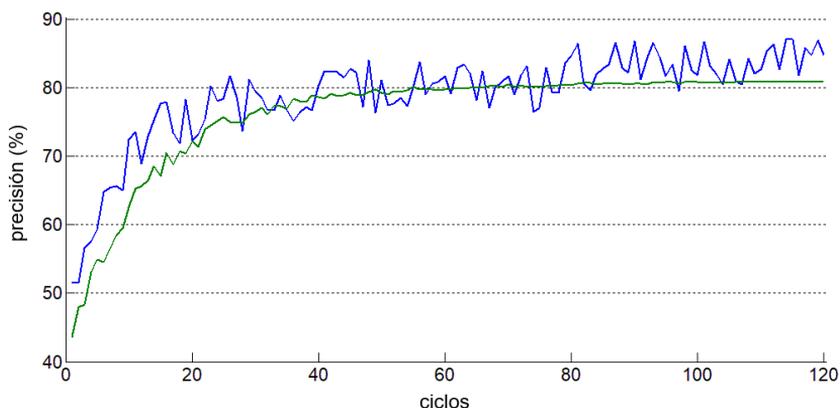
- PC de escritorio convencional, con procesador Intel® Core i7-7700 (Quad Core 3,6 GHz), RAM 16 gb ddr4, GPU NVidia Quadro K1200 4 gb
- Librerías de procesamiento de imágenes y de redes neuronales en lenguaje Python:
  - Keras
  - TensorFlow
- Los parámetros de la red se eligieron basándose en el trabajo de Lakhani et al:
  - 120 ciclos (*epochs*)
  - Ritmo de aprendizaje de 0,001
  - Optimización mediante descenso de gradiente estocástico (*stochastic gradient descend*)
  - Lotes de 64 imágenes por *epoch*

## Materiales y métodos: Medición de performance

- Medición del porcentaje de precisión:
  - Cantidad de placas bien clasificadas sobre el total de placas analizadas en el grupo de validación
- Porcentaje de falsos positivos.
- Porcentaje de falsos negativos.
  
- Obtención de imágenes de Deep Dreaming:
  - Intento visual de comprender las características aprendidas por la red al definir la presencia o ausencia de neumonía.

## Resultados

- Tiempo de entrenamiento aproximado: 5 horas hasta alcanzar la convergencia.

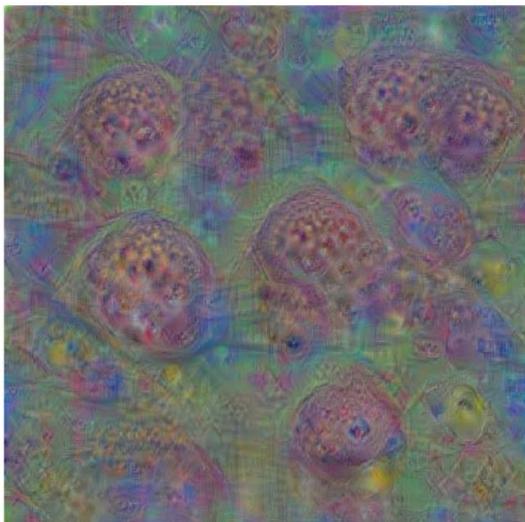
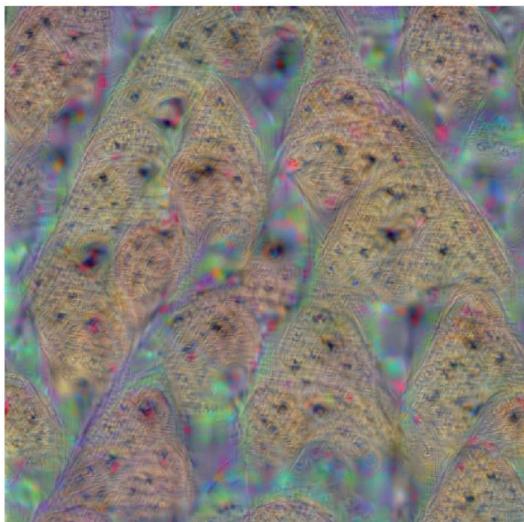


Gráfica de la curva de aprendizaje de una red neuronal donde se puede ver el aumento progresivo de la precisión en el conjunto de entrenamiento (azul) y de validación (verde). Alrededor del ciclo 80 se perciben indicios de sobre-entrenamiento

- P vs Placas NP:
  - Precisión de  $65 \pm 5\%$
  - 13% de falsos positivos (promedio)
  - 22% de falsos negativos (promedio)
- P vs NF:
  - Precisión de  $83 \pm 7\%$
  - 7% de falsos positivos (promedio)
  - 9% de falsos negativos (promedio)

## Resultados

- Imágenes de Deep Dreaming:



Representación del aprendizaje de las CNNs que clasifican (izq) placas con neumonía vs placas sin hallazgos y (der) placas con neumonía vs placas sin neumonía.

## Discusión

- El esquema de *transfer learning* mediante CNNs aplicado a la detección de neumonía mostró un moderado grado de clasificación para diferenciar placas con presencia de neumonía.
- La clasificación mejoró sustancialmente al ser comparados con pacientes sin patologías.
- Posiblemente, la detección automatizada se ve dificultada por el grado de interrelación entre la neumonía y otras patologías (imágenes similares).

## Discusión

- Aunque Rajpurkar et muestran una red con una precisión del 77% para neumonía, los no son comparables con los nuestros:
  - Área bajo la curva ROC (AUC) vs media y desvío en 10 experimentos.
- Además, la red de Rajpurkar et al puede detectar simultáneamente entre 14 patologías diferentes:
  - Nuestro resultado del 83% se consiguió únicamente para P vs NF.
- La imagen de Deep Dreaming muestra que la red ha aprendido a enfocarse en los pulmones, y en la detección de regiones localizadas y de tamaño variable, posiblemente asociadas a la patología a identificar.

## Conclusión

- Mediante un esquema de *transfer learning*, la red AlexNet pudo ser reentrenada en una base de datos libre de 50.000 placas de tórax para la detección de neumonía.
- Los resultados obtenidos son alentadores, aunque aún hay cuestiones para resolver en lo que se refiere a:
  - Confección de bases de datos locales más representativas de nuestra población.
  - Aspectos tecnológicos para mejorar la capacidad diagnóstica de la red:
    - Aumentar la cantidad de capas (profundidad de red)
    - Trabajar con mayor cantidad de imágenes.
    - Trabajar con redes que admitan imágenes de mayor tamaño.

## Bibliografía

- Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights Imaging 2018; Disponible en <https://doi.org/10.1007/s13244-018-0639-9>.
- Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv 2017; Disponible en <https://arxiv.org/abs/1711.05225>.
- Lakhani P, Sundaram B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology 2017; 284(2):574-582.
- 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. arXiv 2017; Disponible en <https://arxiv.org/abs/1705.02315>.
- Buda M, Maki A, Mazurowski MA. A systematic study of the class imbalance problem in convolutional neural networks. arXiv 2017; Disponible en <https://arxiv.org/abs/1710.05381>.
- Simonyan K, Vedaldi A, Zisserman A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. arXiv 2014; Disponible en <https://arxiv.org/abs/1312.6034>.
- Cerda J, Cifuentes L. Using ROC curves in clinical investigation. Theoretical and practical issues. Rev Chil Infect 2012; 29(2): 138-141.