

# Detección De Neumonía Utilizando Redes Convolucionales Y Transfer Learning

**América Victoria Ramírez Cámara**

*Facultad De Ciencias Físico Matemáticas, Universidad Autónoma De Nuevo León, 2022, Nuevo León, México  
america.ramirezcm@uanl.edu.mx*

**Resumen:** La neumonía es una enfermedad común que afecta a millones de personas en el mundo todos los años. Pueden causarla microbios llamados bacterias, virus y hongos. El diagnóstico se basa en los síntomas y el examen físico. Para confirmar el diagnóstico son utilizados las radiografías de tórax, los análisis de sangre y el cultivo del esputo. Hoy en día, existen diversos algoritmos de aprendizaje profundo para el análisis de imágenes médicas. Entre ellos las redes neuronales convolucionales y el transfer learning. En este proyecto se trabajó con una red neuronal compuesta por varias capas convolucionales y dos modelos de transfer learning: Inception Net y Residual Net para la clasificación de radiografías de tórax en neumonía o normal, dando como mejor resultado el modelo de Inception Net con una precisión de 0.86, en comparación con los otros dos modelos con una precisión mucho más baja.

## 1. Introducción

La neumonía es una condición inflamatoria del pulmón que afecta principalmente a los pequeños sacos de aire conocidos como alvéolos. Los síntomas suelen incluir alguna combinación de tos productiva o seca, dolor en el pecho, fiebre y dificultad para respirar. La gravedad de la afección es variable. La neumonía generalmente es causada por infección con virus o bacterias y menos comúnmente por otros microorganismos, ciertos medicamentos o afecciones como enfermedades autoinmunes. Los factores de riesgo incluyen fibrosis quística, enfermedad pulmonar obstructiva crónica (EPOC), asma, diabetes, insuficiencia cardíaca, antecedentes de tabaquismo, mala capacidad para toser, como seguir un accidente cerebrovascular y un sistema inmunitario débil. El diagnóstico a menudo se basa en los síntomas y el examen físico. La radiografía de tórax, los análisis de sangre y el cultivo del esputo pueden ayudar a confirmar el diagnóstico.

Las radiografías de tórax que se utilizan para diagnosticar la neumonía necesitan radioterapeutas expertos para su evaluación, lo que requiere mucho tiempo, es costoso para los hospitales y no necesariamente accesible para los usuarios necesitados. Por lo tanto, el desarrollo de un sistema automático para detectar la neumonía sería beneficioso para tratar la enfermedad sin demora, especialmente en las zonas remotas, y reducir costos en los hospitales.

Debido al éxito de los algoritmos de aprendizaje profundo en el análisis de imágenes médicas, las redes neuronales convolucionales (CNN) han ganado mucha atención para la clasificación de enfermedades. Además, las características aprendidas por modelos CNN previamente entrenados en conjuntos de datos a gran escala son muy útiles en las tareas de clasificación de imágenes.

Para este proyecto, se realizó la clasificación de radiografías de tórax en neumonía o normal utilizando tres modelos de redes neuronales convolucionales (CNN), el primero es una red neuronal compuesta por varias capas convolucionales, el segundo y el tercero son dos modelos de transfer learning: Inception Net y Residual Net.

## 2. Antecedentes

1. Varshni Dimpy et al., 2019. Pneumonia Detection Using CNN based Feature Extraction. <https://ieeexplore.ieee.org/abstract/document/8869364>. Es un artículo de investigación publicado por Varshni Dimpy et al., en el cual se emplearon modelos de Deep Learning (DL), en particular las Redes Neuronales Convolucionales (CNN), que revelaron su potencial para extraer características útiles en tareas de clasificación de imágenes. La disponibilidad de modelos de CNN con formación previa como AlexNet, VGGNet, Xception, ResNet y DenseNet constituye una gran ayuda en el procedimiento de extracción de características significativas. Por lo que ellos emplearon estos modelos mencionados. Observaron el desempeño de varios modelos de CNN pre entrenados junto con clasificadores distintos resultando como los mejores: DenseNet-169 para la etapa de extracción de características y SVM para la etapa de clasificación.

2. V. Sirish Kaushik et al., 2020. Pneumonia Detection Using Convolutional Neural Networks (CNNs). [https://www.researchgate.net/publication/340961287\\_Pneumonia\\_Detection\\_Using\\_Convolutional\\_Neural\\_Networks\\_CNNs](https://www.researchgate.net/publication/340961287_Pneumonia_Detection_Using_Convolutional_Neural_Networks_CNNs). Es un artículo de investigación publicado por V. Sirish Kaushik et al., en el cual crearon los modelos de CNN desde cero y fueron entrenados en imágenes de rayos X de tórax (neumonía) en Kaggle. La biblioteca de la red neuronal de Keras con TensorFlow se utilizó para implementar los modelos. El dataset consta de 5216 imágenes de entrenamiento, 624 imágenes de prueba y 16 imágenes de validación. El aumento de datos se aplicó para lograr mejores resultados del conjunto de datos. Los cuatro modelos se entrenaron en el conjunto de datos de entrenamiento, cada uno con un número diferente de capas convolucionales. Cada modelo fue entrenado para 20 épocas, con entrenamiento y probando tamaños de lote de 32 y 1, respectivamente. De sus cuatro modelos, el modelo 3 y 4 obtuvieron los mejores resultados.
3. Alhassan Mabrouk et al., 2022. Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks. <https://www.mdpi.com/2076-3417/12/13/6448>. En este trabajo de investigación, los autores utilizaron tres conocidos modelos (DenseNet169, MobileNetV2 y Vision Transformer) previamente preparados utilizando la base de datos ImageNet. Estos modelos se entrenan en el conjunto de datos de rayos X de tórax mediante el ajuste fino. Finalmente, los resultados se obtienen combinando las características extraídas de estos tres modelos durante la fase experimental. El enfoque propuesto supera a otros métodos de vanguardia existentes y obtiene una precisión del 93,91 % y una puntuación F1 del 93,88 % en la fase de prueba.

### 3. Descripción De Datos

El conjunto de datos está organizado en 3 carpetas (train, test, val) y contiene subcarpetas para cada categoría de imagen (Pneumonia/Normal). Hay 5863 imágenes de rayos X (JPEG) y 2 categorías (Pneumonia/Normal). Se seleccionaron imágenes de rayos X de tórax (anterior-posterior) de pacientes pediátricos de uno a cinco años del Centro Médico de Mujeres y Niños de Guangzhou, Guangzhou. Todas las imágenes de rayos X de tórax se realizaron como parte de la atención clínica de rutina de los pacientes. Para el análisis de las imágenes de rayos X de tórax, todas las radiografías de tórax se examinaron inicialmente para el control de calidad mediante la eliminación de todas las exploraciones de baja calidad o ilegibles. Los diagnósticos de las imágenes fueron calificados por dos médicos expertos antes de ser aprobados para entrenar el sistema de IA. Con el fin de dar cuenta de cualquier error de calificación, el conjunto de evaluación también fue verificado por un tercer experto.

### 4. Metodología

#### 4.1. Carga y División de Datos

En esta sección se cargaron los datos de la carpeta de entrenamiento (train) y validación (val) y se realizó la división de estos en un ratio de 80:20 ya que la cantidad de imágenes de entrenamiento son de 5216 (neumonía: 3875, normal: 1341) y de validación 16 (neumonía: 8, normal: 8), era mucha la diferencia por lo que se realizó este proceso para un alto rendimiento en los modelos y la obtención de buenos resultados. Después de la división, el total de imágenes de entrenamiento es de 4185 (neumonía: 3105, normal: 1080) y de validación 1047 (neumonía: 778, normal: 269).

#### 4.2. Aumento de Datos

Utilizando ImageDataGenerator, se aumentó el conjunto de datos para que el aprendizaje sea difícil para un modelo dado, para que el modelo pueda aprender características diferentes y difíciles desde el dataset de entrenamiento que nos ayuda a obtener una buena precisión en el dataset de prueba (test).

#### 4.3. Corrección para el desequilibrio de datos

Luego de analizar el conjunto de datos de entrenamiento después de la división descrita en la primera subsección, había muchas imágenes de neumonía (positivo) en comparación con las imágenes normales (negativo). Esto mostraba un desequilibrio en los datos. Por lo que se realizó una corrección. El peso para la clase 0 (normal) es mucho más alto que el peso para la clase 1 (neumonía). Debido a que hay imágenes menos normales, cada imagen normal se ponderará más para equilibrar los datos, ya que la CNN funciona mejor cuando los datos de entrenamiento están equilibrados. Se utilizaron las siguientes formulas para calcular los pesos de ambas clases, con el fin de ser utilizados posteriormente en el entrenamiento del modelo.

$$WeightForNormal = \frac{\frac{1}{CountNormal} \times TotalTrainImages}{2.0}$$

$$WeightForPneumonia = \frac{1}{\overline{CountPneumonia}} \times \frac{TotalTrainImages}{2.0}$$

#### 4.4. Entrenamiento de los Modelos

Para realizar la clasificación de nuestros datos se entrenaron tres modelos diferentes para comparar su rendimiento y los resultados obtenidos en cada uno. Los dos últimos modelos de transfer learning: Inception Net y Residual Net fueron utilizados por Varshni Dimpy et al. en el 2019 donde describen en su artículo titulado “Pneumonia Detection Using CNN based Feature Extraction,” que estos modelos constituyen una gran ayuda en el procedimiento de extracción de características significativas; además de obtener buenos resultados en exactitud en su proyecto. Esta es la razón por la cual se eligieron estos dos modelos de transfer learning.

##### 4.4.1. Modelo CNN

Este es un modelo que consta de varias capas convolucionales, y el cual fue entrenado utilizando las funciones de activación relu y sigmoid; y el optimizador rmsprop junto con la función de pérdida binary crossentropy. En la figura 1, se puede observar el diseño de dicho modelo.

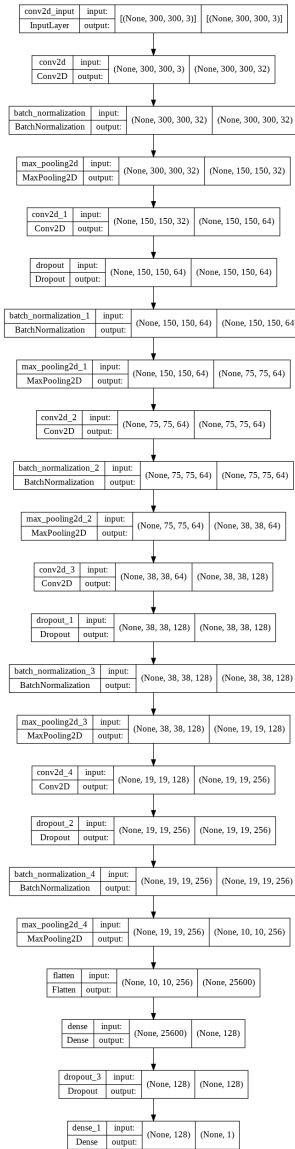


Figura 1. Diseño del modelo CNN.

#### 4.4.2. Modelo Inception Net

Este es un modelo de transfer learning, y el cual fue entrenado utilizando las función de activación sigmoid; y el optimizador rmsprop junto con la función de pérdida binary crossentropy. En la figura 2, se puede observar el diseño de dicho modelo.

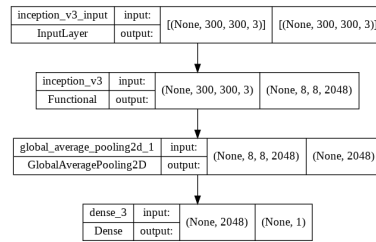


Figura 2. Diseño del modelo Inception Net.

#### 4.4.3. Modelo Residual Net

Este es un segundo modelo de transfer learning, y el cual fue entrenado utilizando las función de activación sigmoid; y el optimizador rmsprop junto con la función de pérdida binary crossentropy. En la figura 3, se puede observar el diseño de dicho modelo.

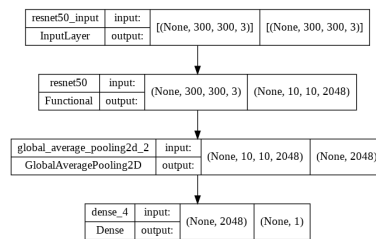


Figura 3. Diseño del modelo Residual Net.

Por último, después del entrenamiento de cada modelo, se realizaron algunos gráficos para visualizar el rendimiento, una predicción y evaluación en el dataset de prueba. En la sección de resultados se observan estos detalles.

## 5. Resultados

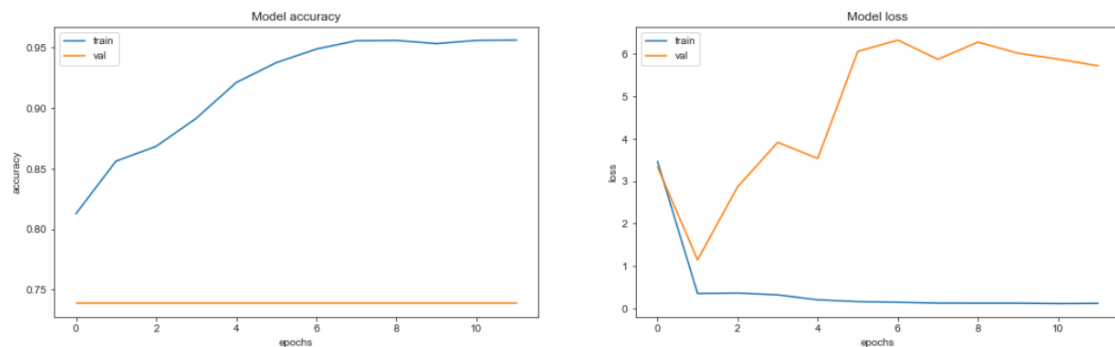


Figura 4. Rendimiento del modelo CNN visualizando las métricas de exactitud y pérdida.

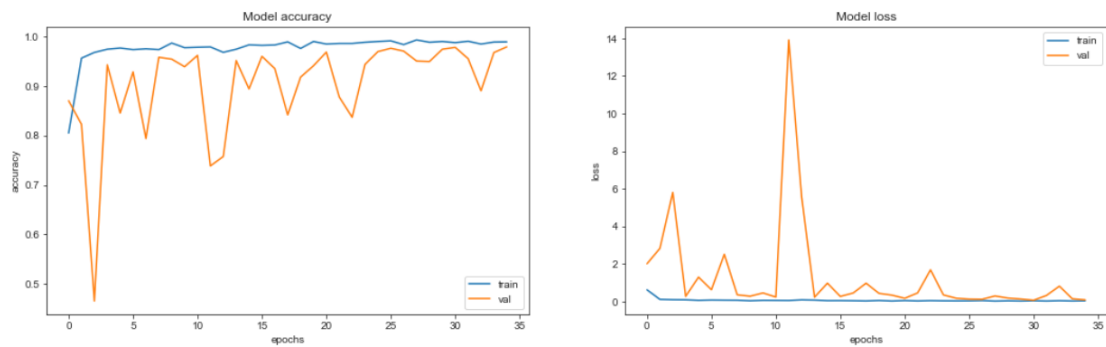


Figura 5. Rendimiento del modelo Inception Net visualizando las métricas de exactitud y pérdida.

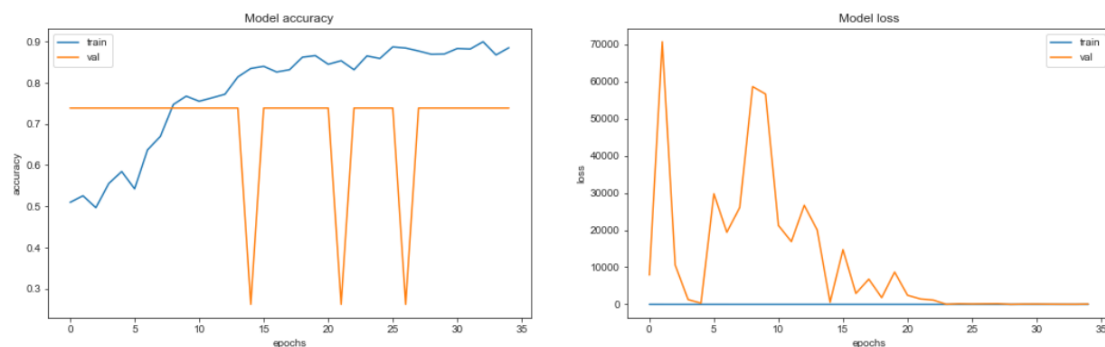


Figura 6. Rendimiento del modelo Residual Net visualizando las métricas de exactitud y pérdida.

Tabla 1. Rendimiento de los Modelos

Modelo	Épocas	Exactitud	Pérdida
Modelo CNN	12	0.625	9.311
Modelo Inception Net	35	0.868	0.915
Modelo Residual Net	35	0.625	131.042

## 6. Discusión

Para el modelo CNN, en la figura 4 se puede observar la exactitud de los datos de entrenamiento (train) contra los datos de validación (val). La exactitud de los datos de validación se mantuvieron constantes durante las 12 épocas, a comparación de los datos de entrenamiento donde esta métrica fue aumentando. Se esperaba que fueran similares entre ambos, pero se observa una gran diferencia. También cabe recalcar que la pérdida fue aumentando en los datos de validación, contrario a los datos de entrenamiento que se observa de manera descendente al principio y después se mantiene constante.

Para el modelo Inception Net, en la figura 5 se puede observar la exactitud de los datos de entrenamiento (train) contra los datos de validación (val). La exactitud de los datos de validación se mantuvieron de forma estacionaria durante las 35 épocas, ascendiendo y descendiendo, a comparación de los datos de entrenamiento donde esta métrica se mantuvo casi constante. También cabe recalcar que la pérdida tuvo algunos ascensos en los datos de validación, contrario a los datos de entrenamiento que se observa de manera casi constante.

Para el modelo Residual Net, en la figura 6 se puede observar la exactitud de los datos de entrenamiento (train) contra los datos de validación (val). La exactitud de los datos de validación tuvo ascensos, descensos y valores constantes durante las 35 épocas, se comportó de una manera poco común, a comparación de los datos de entrenamiento donde esta métrica fue aumentando. Se esperaba que fueran similares entre ambos, pero se observa una gran diferencia. También cabe recalcar que la pérdida tuvo ascensos y descensos en los datos de validación, contrario a los datos de entrenamiento que se mantiene constante.

En la tabla 1, se puede observar el total de épocas asignadas a cada modelo, junto con su exactitud y pérdida, siendo evaluados para los datos de pruebas. Los mejores resultados se obtuvieron con el modelo Inception Net con un 0.868 de exactitud y una pérdida baja de 0.915.

En el artículo de Varshni Dimpy et al. en el año 2019 titulado “Pneumonia Detection Using CNN based Feature Extraction.”, se trabajó con los mismos modelos de transfer learning, sin embargo, Inception Net tuvo mejor exactitud en mi metodología comparando contra un 0.70 alcanzado en el estudio descrito en este artículo. El aumento de datos es una parte importante que cabe resaltar desde que nos ayuda a obtener una buena exactitud.

## 7. Conclusiones

La presencia de radioterapeutas expertos es la mayor necesidad para diagnosticar adecuadamente cualquier tipo de enfermedad torácica. El objetivo principal de este trabajo es mejorar la aptitud médica en áreas donde la disponibilidad de radioterapeutas es todavía limitada. Este proyecto puede facilitar el diagnóstico temprano de la neumonía para prevenir consecuencias adversas (incluyendo la muerte) en áreas remotas. El desarrollo de algoritmos en este ámbito puede ser muy beneficioso para proporcionar mejores servicios de atención de la salud. Observamos tres modelos, los cuales mostraron resultados muy diferentes, siendo el mejor: Inception Net con una exactitud de 0.86. Con la serie de experimentos realizados, el objetivo es proporcionar el modelo de CNN preponderante para el trabajo futuro en el campo de la investigación similar. Este proyecto conducirá al desarrollo de mejores algoritmos para la detección de la neumonía en el futuro previsible.

## Referencias

1. MATHUR, M. (2020). *Pneumonia Detection using CNN(92.6% Accuracy)* [Online]. Disponible: <https://www.kaggle.com/code/madz2000/pneumonia-detection-using-cnn-92-6-accuracy/data>
2. DIMPY, V. ET AL. (2019). *Pneumonia Detection Using CNN based Feature Extraction* [Online]. Disponible: <https://ieeexplore.ieee.org/abstract/document/8869364>
3. KAUSHIK, V. ET AL. (2020). *Pneumonia Detection Using Convolutional Neural Networks (CNNs)* [Online]. Disponible: [https://www.researchgate.net/publication/340961287\\_Pneumonia\\_Detection\\_Using\\_Convolutional\\_Neural\\_Networks\\_CNNs](https://www.researchgate.net/publication/340961287_Pneumonia_Detection_Using_Convolutional_Neural_Networks_CNNs)
4. MABROUK, A. ET AL. (2022). *Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks* [Online]. Disponible: <https://www.mdpi.com/2076-3417/12/13/6448>
5. SHEKHAR, C. (2020). *VGG(91%), InceptionNet(90%) to classify pneumonia* [Online]. Disponible: <https://www.kaggle.com/code/csk03012/vgg-91-inceptionnet-90-to-classify-pneumonia/notebook>
6. GITHUB (2022). *ProcesamientoDeDatos* [Online]. Disponible: <https://github.com/vickymz24/ProcesamientoDeDatos/blob/main/ProyectoFinal.ipynb>