



U N I V E R S I D A D
Panamericana

FACULTAD DE INGENIERÍA

**“SEGMENTACIÓN AUTOMÁTICO DE NEUMOTÓRAX EN
IMÁGENES DE RAYOS X”**

TESIS

P R E S E N T A
SEBASTIÁN GARCÍA AGUIRRE

PARA OBTENER EL GRADO DE
MAESTRÍA EN CIENCIAS

CON RECONOCIMIENTO DE VALIDEZ OFICIAL DE ESTUDIOS DE LA
SECRETARÍA DE EDUCACIÓN PÚBLICA, SEGÚN ACUERDO CON EL No. 2007574
DE FECHA 29 de junio de 2007

DIRECTORES DE TESIS
DRA. CLAUDIA NALLELY SÁNCHEZ GÓMEZ
DR. MARIANO JOSÉ JUAN RIVERA MERAZ

AGUASCALIENTES, AGS., julio de 2022

Biblioteca Aguascalientes



Aguascalientes, Ags., 4 de julio de 2022

LIBERACIÓN DE TESIS

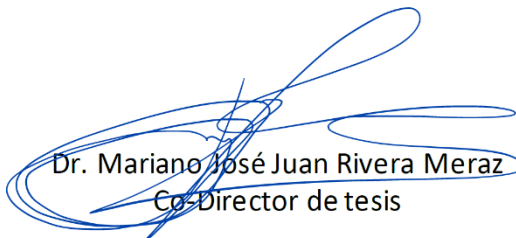
Por medio de la presente, certificamos en nuestra calidad de directores de tesis, que el trabajo de SEBASTIÁN GARCÍA AGUIRRE que lleva como título: SEGMENTACIÓN AUTOMÁTICO DE NEUMOTÓRAX EN IMÁGENES DE RAYOS X cumple con los requisitos establecidos por el reglamento vigente de la Facultad de Ingeniería para presentarse en Examen de Titulación del programa de Maestría en Ciencias.

De resultar aprobado, podrá efectuar el trámite para la obtención del grado de Maestría en Ciencias.

Atentamente,



Dra. Claudia Nallely Sánchez Gómez
Directora de tesis



Dr. Mariano José Juan Rivera Meraz
Co-Director de tesis

Josemaría Escrivá de Balaguer 101, Fracc. Villas Bonaterra, Aguascalientes 20296, México. Tel. +52 (449)9106200
www.up.edu.mx

Biblioteca Aguascalientes

Agradecimientos

Quiero agradecer el apoyo para la realización de esta tesis a:

Mi asesora de tesis, Claudia Sánchez, por el apoyo, la motivación, la paciencia en todo este nuevo proceso de mi vida profesional. Siempre estaba dispuesta a ayudarme desde mi formación universitaria, con mis problemas tanto personales como los que se presentaron en esta tesis.

Al profesor Mariano Rivera, por su alto conocimiento del área de aprendizaje de máquina. Siempre aportaba múltiples soluciones a problemas que se tenían del área.

A la Universidad Panamericana Aguascalientes, por la preparación profesional que me ofreció y los conocimientos para poder utilizar nuevas tecnologías para poder realizar esta tesis, y el hardware utilizado para el aprendizaje del modelo.

A mis padres por el ánimo y el apoyo para poder concluir con todos mis estudios.

A mis hermanos, hermana y nana por todo su amor y complicidad.

A mis amigos por la felicidad que nos aportamos y los ánimos que tenemos en nuestro entorno.

Biblioteca Aguascalientes

Índice general

AGRADECIMIENTOS	5
ÍNDICE GENERAL	7
LISTA DE FIGURAS	8
LISTA DE TABLAS	9
GLOSARIO	10
RESUMEN.....	11
ABSTRACT	12
1. Introducción	15
<i>Justificación.....</i>	17
<i>Objetivos.....</i>	17
<i>Hipótesis</i>	17
2. Fundamentos teóricos.....	19
2.1 <i>Neurona artificial.....</i>	19
2.2 <i>Redes Neuronales Artificiales.....</i>	21
2.3 <i>Redes Neuronales Convolucionales</i>	22
2.4 <i>Redes Neuronales de Arquitectura Codificador - Decodificador.....</i>	24
2.5 <i>Estado del Arte</i>	26
3. Metodología	29
3.1 <i>Conjunto de Imágenes</i>	29
3.2 <i>Conjunto de datos para entrenar la red neuronal para segmentación.....</i>	30
3.3 <i>Segmentación de Imágenes de Neumotórax.....</i>	30
3.4 <i>Arquitectura de Red Neuronal Codificador - Decodificador</i>	31
4. Experimentos y Resultados	36
4.1 <i>Equipo de cómputo.....</i>	36
4.2 <i>Convergencia de la red neuronal.....</i>	36
4.3 <i>Resultados en términos del coeficiente Jaccard.....</i>	37
4.4 <i>Resultados en imágenes</i>	38
5. Conclusión	44
5.1 <i>Conclusión</i>	44
5.2 <i>Trabajo a futuro</i>	44
6. Bibliografía	46

Lista de figuras

Figura 1.1 Imagen de Rayos X con neumotórax.	15
Figura 2.1 Representación de una neurona artificial.....	19
Figura 2.2 Representación de una red neuronal artificial	21
Figura 2.3 Ejemplo de Convolucion	22
Figura 2.4 Ejemplo de Pooling.	23
Figura 2.5 Ejemplo de arquitectura de una red neuronal convolucional	23
Figura 2.6 Ejemplo de red neuronal codificador-decodificador.....	24
Figura 2.7 Ejemplo de UpSampling.....	25
Figura 2.8 Ejemplo de Concatenación	25
Figura 3.1 Ejemplo de formato de representación RLE. Imagen propia.....	30
Figura 3.2 Imagen de Rayos X con neumotórax.	31
Figura 3.3 Arquitectura de la red neuronal utilizada para la segmentación de Neumotórax.....	32
Figura 3.4 Arquitectura de la red neuronal utilizada para la segmentación de Neumotórax con detalle	33
Figura 3.5 Inicio de la arquitectura de la red neuronal utilizada para la segmentación de Neumotórax con detalle	34
Figura 3.6 Fin de la arquitectura de red neuronal utilizada para la segmentación de Neumotórax con detalle	34
Figura 4.1 Convergencia del entrenamiento de la red neuronal.....	36
Figura 4.2 Operaciones de intersección y unión en conjuntos.....	37
Figura 4.3 Macro-Jaccard de los conjuntos de imágenes de entrenamiento, validación y prueba.	38
Figura 4.4 Resultados de la segmentación de neumotórax.	42

Lista de tablas

Tabla 4.1 Promedio de Macro-Jaccard de los conjuntos de imágenes de entrenamiento, validación y prueba.....	38
---	----

Biblioteca Aguascalientes

Glosario

Imagen de rayos X: Imágenes binarias creadas por la radiación de radiación electromagnética.

Neumotórax: Afección cuando el aire entra entre la pleura y el pulmón.

Red Neuronal Artificial: Red interconectada de nodos, inspirado en un cerebro humano

Redes Neuronales Codificador-Decodificador: red neuronal especializada para segmentación de imágenes.

Red Neuronal Convolutiva (CNN): Modelo de red neuronal utilizada principalmente para imágenes basado en arquitectura convolutiva

Segmentación: Dividir una imagen en regiones con características similares.

SIIM: Sociedad de informática Imágenes en Medicina.

Biblioteca Aguascalientes

Resumen

El neumotórax es un padecimiento que afecta al pulmón. Es la presencia de aire entre el pulmón y la pleura, dificultando la expansión del pulmón y la respiración del paciente. El objetivo de la tesis es presentar una metodología para la segmentación de neumotórax de forma automática utilizando redes neuronales de aprendizaje profundo a través de imágenes de rayos X. Una red neuronal artificial es un conjunto de neuronas artificiales, que simulan el funcionamiento de un cerebro humano, pudiendo así encontrar patrones complicados y predecir una respuesta. La base de datos de imágenes de rayos X se tomó del concurso SIIM-ACR Pneumothorax Segmentation propuesto por la SIIM (Society for Imaging Informatics in Medicine) en la plataforma Kaggle, cuyo objetivo es desarrollar un sistema de Inteligencia Artificial capaz de segmentar automáticamente la zona de neumotórax en imágenes de rayos X. Contiene 12,089 imágenes de rayos X de la zona pulmonar. En este proyecto solo se usaron las imágenes que tenían neumotórax. La red neuronal propuesta es una red neuronal codificador-decodificador que tiene como principal área de uso los problemas de segmentación. El modelo propuesto obtuvo un resultado en promedio de 0.6072 para las imágenes de prueba utilizando como métrica el coeficiente Macro-Jaccard. Es una medida que va del 0 al 1, siendo 1 el mejor. Nuestros resultados presentan una solución adecuada al problema de segmentación de neumotórax. Cabe mencionar que esta tecnología también puede ser aplicada a otras áreas de imágenes médicas para determinar las zonas de diferentes enfermedades.

Palabras clave: Neumotórax, Aprendizaje profundo, Segmentación, Rayos X, Imágenes Médicas.

Abstract

Pneumothorax is a condition that affects the lungs. It is the air between the lung and the pleura, hindering the expansion of the lung and affecting the patient's breathing. The thesis aims to present a methodology for the segmentation of pneumothorax automatically using deep learning neural networks through X-ray images. An artificial neural network is a set of artificial neurons that simulate a human brain's functioning, thus being able to find complicated patterns and predict an answer. The X-ray image database was taken from the SIIM-ACR Pneumothorax Segmentation contest proposed by the SIIM (Society for Imaging Informatics in Medicine) on the Kaggle platform, whose objective is to develop an Artificial Intelligence system capable of automatically segmenting the area of pneumothorax in X-ray images. This dataset contains 12,089 X-ray images of the lung area. Only the images that had pneumothorax were used in this project. The neural network proposal is a neural network encoder-decoder whose main area of use is segmentation problems. The proposed model obtained an average result of 0.6072 for the test images using the Macro-Jaccard coefficient as a metric. It goes from 0 to 1, with 1 being the best result. Our results present an adequate solution to the pneumothorax segmentation problem. It is worth mentioning that this technology can also be applied to other areas of medical imaging to determine the areas of different diseases.

Keywords: Pneumothorax, Deep Learning, Segmentation, X Ray, Medical Images.

Biblioteca Aguascalientes

Capítulo 1

Introducción

Biblioteca Aguascalientes

1. Introducción

El neumotórax es una afección potencialmente mortal caracterizada por la presencia anormal de aire en el espacio pleural (Li et al. 2019). Esto ocurre cuando el aire se filtra entre los pulmones y la cavidad torácica afectando a la capacidad de los pulmones para expandirse al tomar aire o incluso puede llegar a colapsar los pulmones. El neumotórax puede ocurrir por alguna lesión penetrante en el pecho, cirugías o daño por algún problema pulmonar subyacente. El tratamiento consiste en introducir una aguja entre las costillas para liberar el excedente de aire.

El neumotórax se detecta, en las imágenes de rayos X, mediante la detección de la reducción de tamaño del pulmón (analizando el contorno del pulmón). Principalmente se localiza en la parte superior, lateral e inferior, debido a que el material atrapado principalmente está conformado por aire. Esto también se puede ver, mediante la falta de los alvéolos (ausencia de marcas/vasos pulmonares).

En la Figura 1.1 se puede observar un ejemplo de una imagen de rayos X tomada de la zona de pulmón. En la imagen se muestra la necesidad de representar la zona de neumotórax con una imagen adicional que contiene la región donde se presenta.

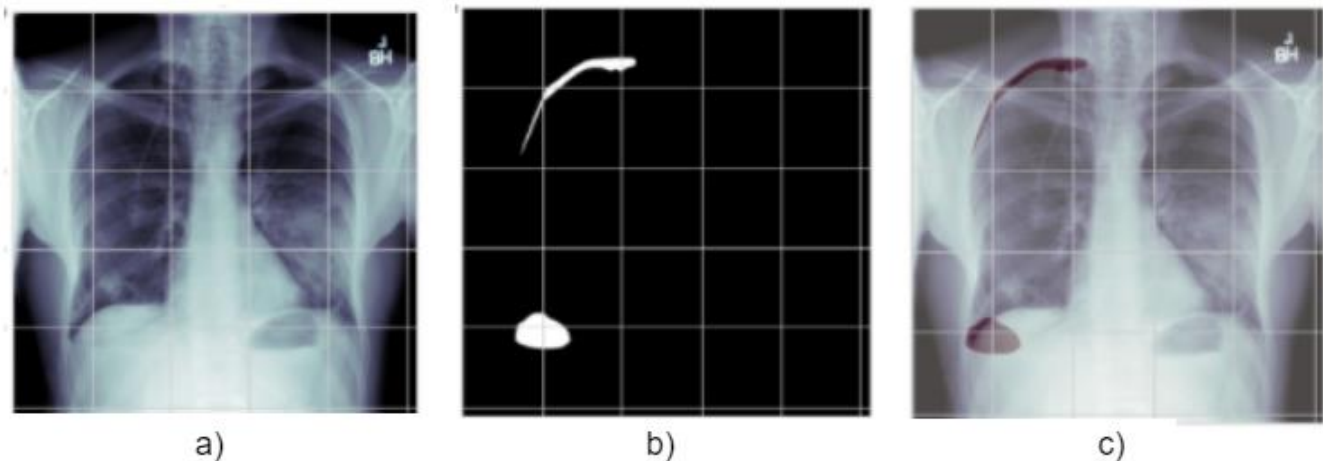


Figura 1.1 Imagen de Rayos X con neumotórax.

- a) Imagen original en rayos X. b) Imagen que representa la máscara indicando la zona del neumotórax.
c) Combinación de a) y b).

La tecnología avanza muy rápidamente y hay una oportunidad muy importante para aplicarla en diagnósticos clínicos de múltiples enfermedades que pueden ser diagnosticadas en base a programas computacionales. La aplicación de la Inteligencia Artificial ya es una realidad aún y cuando este término fue adoptado en 1956, hoy en día se ha vuelto más popular debido a que ya se cuenta con una mayor cantidad de datos que nos permiten tener mayor información sobre un problema específico, así como el

desarrollo de algoritmos avanzados, capacidad de almacenaje y computadoras más potentes que pueden procesar estos datos con mayor facilidad y rapidez.

La Inteligencia Artificial hace posible que las computadoras puedan aprender a realizar tareas específicas procesando grandes cantidades de información histórica y reconociendo patrones en los datos. Las redes neuronales artificiales son modelos computacionales que simulan el funcionamiento del cerebro humano, pudiendo resolver problemas muy complejos. Estos funcionan con nodos simulando neuronas, las cuales están conectadas entre sí para recibir y enviar señales a otros nodos. Las redes neuronales son flexibles y adaptables, logrando resolver la mayoría de los problemas y actualmente están presentes en infinidad de objetos del día a día. En los últimos años, avances en las capacidades computacionales han logrado desarrollar modelos de redes neuronales profundas o también conocido simplemente como Aprendizaje Profundo, con una gran cantidad de capas de neuronas, que han permitido lograr avances significativos en diversas áreas del conocimiento como los ejemplos siguientes. Vehículos autónomos (Grigorescu et al. 2020). Diseño molecular (Elton et al. 2019). Procesamiento de rostros, por ejemplo, identificación (Balaban 2015), predicción de rostros al envejecer y rejuvenecer (Antipov et al. 2017). Ciberseguridad (Xin et al. 2018). Clasificación de objetos 3D (Socher et al. 2012). Reconocimiento, detección y segmentación de objetos 3D (Wang 2016)

En el campo de la medicina, las técnicas de la inteligencia artificial como aprendizaje profundo con redes neuronales, se tienen avances significativos, tal que, en algunos casos, los resultados de estos modelos computacionales pueden ser comparados con el criterio humano. A continuación, se presentan algunos ejemplos: Segmentación de tumores cerebrales en imágenes MRI (Işın et al. 2016). Detección de Alzheimer en imágenes MRI (Altinkaya et al. 2020). Detección de Covid-19 (Alazab et al. 2020) Clasificación de cáncer de piel (Brinker et al. 2018).

En este documento se propone una metodología basada en Aprendizaje Profundo para la detección automática de la zona con neumotórax de tal forma que este proceso pueda ser automatizado.

Justificación

El avance de la tecnología permite la automatización de las tareas realizadas por los seres humanos, de tal forma que puedan ser ejecutadas con una mayor rapidez y sobre todo con mayor certeza. Específicamente, en el caso de detección de enfermedades, los médicos desarrollan y mejoran su habilidad de acuerdo con su experiencia y a los casos que se enfrentan. En cambio, un modelo computacional puede aprender utilizando un conjunto de datos bastante grande que puede provenir de diferentes tipos de personas y países, pudiendo así generalizar y agilizar el proceso de entrenamiento.

Hoy en día, el aprendizaje profundo ha generado avances bastantes significativos al campo de la Inteligencia Artificial, es por esto que consideramos es muy importante utilizarlo también para resolver problemas médicos como este en específico que es Segmentación del Neumotórax.

Objetivos

Crear y entrenar una red neuronal convolucional capaz de identificar y segmentar, la presencia de neumotórax en un paciente mediante una imagen médica de rayos X de la zona pulmonar del paciente.

La red neuronal convolucional debe ser capaz de segmentar el área de la imagen de rayos X donde se represente el neumotórax ayudando así a facilitar y agilizar el tratamiento de la enfermedad.

Objetivos particulares:

- Seleccionar una base de datos de imágenes de rayos X de la zona pulmonar de pacientes donde algunos tengan problema de neumotórax y otros no. Adicionalmente, a las imágenes de rayos X, para los pacientes con problema de neumotórax se deben tener las imágenes con las marcas que representan la zona afectada (ver Figura 1a).
- Normalizar los datos obtenidos de la base de datos. Es decir, todas las imágenes deben tener el mismo tamaño.
- Entrenar un modelo de segmentación de la zona de neumotórax por medio de una red neuronal convolucional profunda utilizando imágenes de rayos X como entrenamiento.

Hipótesis

Es posible generar un modelo de red neuronal convolucional que sea capaz de identificar la zona específica de neumotórax a través de imágenes de rayos X.

Capítulo 2

Fundamentos Teóricos

2. Fundamentos teóricos

En este capítulo se explicará a detalle los fundamentos de las redes neuronales artificiales, y las diferentes arquitecturas como redes neuronales convolucionales y redes neuronales codificador-decodificador. Así como su aplicación en problemas de visión por computadora.

2.1 Neurona artificial

Una neurona es una célula del sistema nervioso que se encarga de recibir, procesar y emitir estímulos eléctricos que resultan en una respuesta del cuerpo humano. A su vez, una neurona artificial es la parte más básica de una red neuronal artificial. Las neuronas artificiales son parecidas a las neuronas humanas, ya que tienen interconexiones entre otras neuronas y al procesar las entradas, regresa un resultado.

El proceso matemático de una neurona se puede definir en la Ecuación 2.1 y se describe a continuación (ver Figura 2.1). La neurona recibe una o varias entradas numéricas ($x_1, x_2, x_3, \dots, x_n$), cada una asociada con un peso ($w_1, w_2, w_3, \dots, w_n$). La salida de la neurona será la suma de cada una de las entradas multiplicada por su respectivo peso, y después la salida se procesa a través de una función de activación (F). Esta salida será la entrada para una neurona siguiente. Las entradas numéricas de la red neuronal deben ser pre-procesadas para que tengan valores entre 0 y 1 simulando estímulos eléctricos.

$$y = F(\sum_i w_i, x_i) \tag{2.1}$$

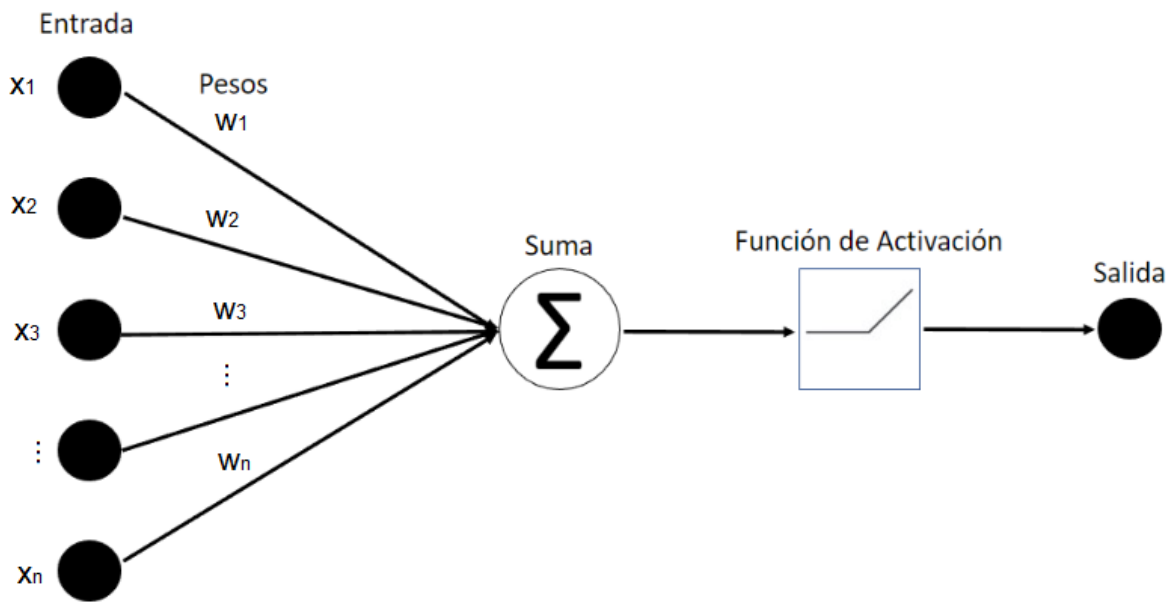
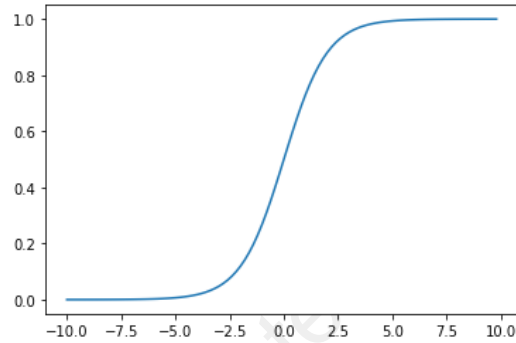


Figura 2.1 Representación de una neurona artificial.

La función de activación tiene por objetivo generar modelos no lineales, además de enviar la información con valores entre 0 y 1. Existen diferentes tipos de funciones de activación, a continuación, se presentan los más utilizados:

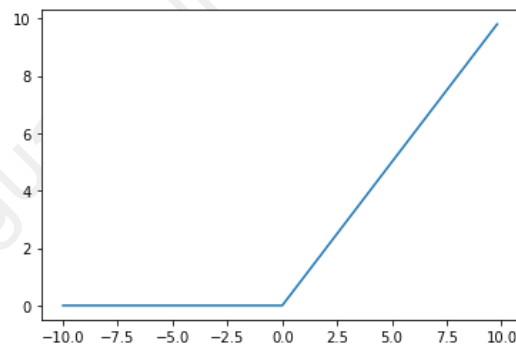
- Función Sigmoide

$$\frac{1}{1 + e^{-x}}$$



- Función ReLu

$$\text{Max}(0, x)$$



2.2 Redes Neuronales Artificiales

Una red neuronal es una interconexión de nodos (neuronas) imitando una estructura cerebral. Con esto se logra “aprender” ciertos patrones en los datos para luego poder predecir o generalizar alrededor de lo que ya se haya aprendido. El término red neuronal tiene sus orígenes en intentos por encontrar representaciones matemáticas del proceso de información en sistemas biológicos (McCulloch and Pitts, 1943). Los nodos o neuronas artificiales se organizan en capas, las cuales son un conjunto de neuronas organizadas para encontrar diferentes patrones. Como se muestra en la Figura 2.2, podemos distinguir principalmente tres áreas donde se encuentran las capas de neuronas:

- Capa de entrada: Es donde las neuronas reciben los datos de entrada.
- Capas ocultas: Son capas que procesan la información mediante operaciones matemáticas. Podrán encontrar patrones para determinar la predicción.
- Capa de salida: El modelo regresa los resultados a través del procesamiento de las capas ocultas.

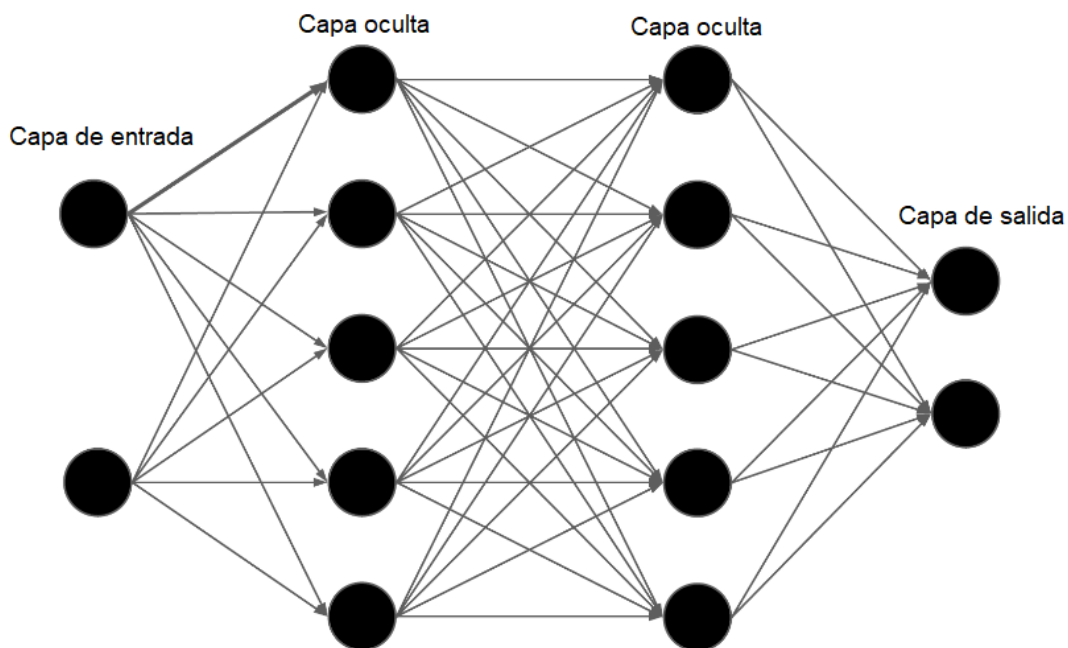


Figura 2.2 Representación de una red neuronal artificial

2.3 Redes Neuronales Convolucionales

Las redes neuronales convolucionales (CNN por sus siglas en inglés) son un tipo de redes neuronales donde las neuronas reaccionan similarmente a campos visuales del cerebro humano. Haciéndolas así muy útiles y efectivas para tareas de visión por computadora, como la categorización o segmentación de imágenes. Se componen de varias capas de neuronas. Estas capas pueden tener diferentes funciones y conexiones. A continuación, se explican las capas más comunes.

Capas de convolución: Este tipo de capa se usa principalmente para procesar las imágenes utilizando su estructura matricial. Usa una máscara de *kernel* que recorre cada píxel de la imagen, y regresa una convolución entre la matriz de la imagen y la matriz de kernel. En la Figura 2.3 se muestra un ejemplo de convolución que se explica a continuación. La máscara de kernel es una matriz de 3x3, por lo tanto, se toman ventanas de píxeles de 3x3 para realizar el cálculo. En la imagen se puede ver la ventana centrada en el píxel de la tercera fila y tercera columna. Cada píxel en la ventana se multiplica por el valor de la máscara de kernel, esto da como resultado otra matriz de tamaño 3x3, de la cual se suman todos sus valores para obtener el valor del píxel en la nueva matriz resultante de la convolución. En la etapa de aprendizaje de la red neuronal, se aprenden los valores de las matrices de kernel.

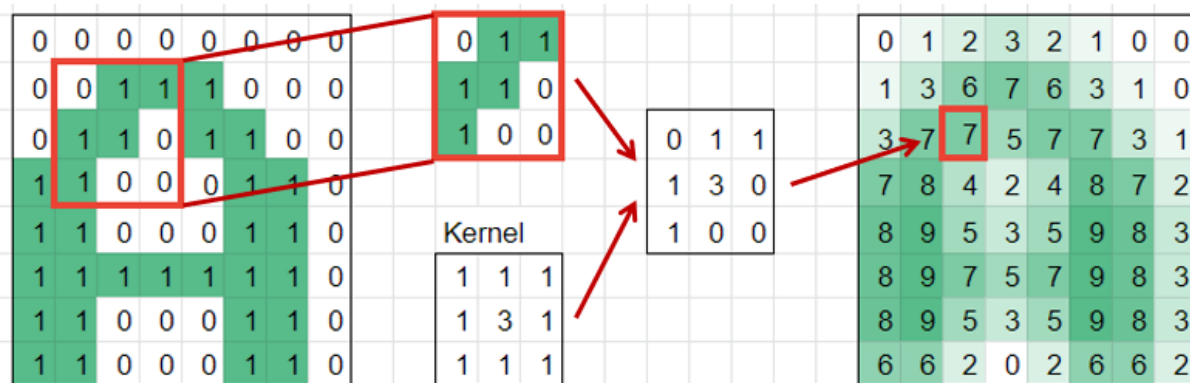


Figura 2.3 Ejemplo de Convolucion

Capa Dropout: La capa de dropout desactiva algunas neuronas en ciertas iteraciones del proceso de entrenamiento. El objetivo es que aleatoriamente se seleccionen los pesos que se aprenden en dicha iteración. El dropout ayuda a prevenir el sobreajuste. Generalmente vienen asociadas a una capa de convolución. El dropout fue inicialmente propuesto por Srivastava et al. (2014).

Capas Pooling: El proceso de pooling consiste en reducir el tamaño de una imagen a la mitad utilizando ventanas. Reduce el tamaño de la entrada en sus dimensiones, tomando el valor máximo en una ventana de tamaño definido, facilitando el procesamiento de la información. La Figura 2.4 muestra un ejemplo.

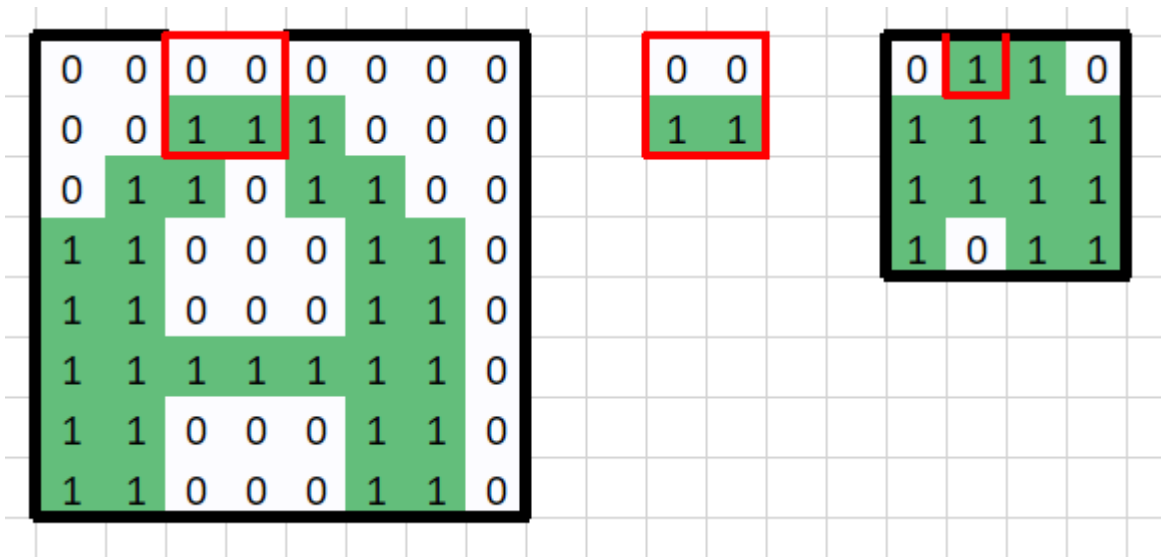


Figura 2.4 Ejemplo de Pooling.

Capas Densas: Este tipo de capa conecta todos sus nodos con todos los nodos de la capa siguiente. Generalmente en una red neuronal convolucional, vienen hasta el final para procesar los resultados de las capas anteriores en una o varias salidas numéricas simples. Se podrían ver como las capas ocultas de la Figura 2.2.

La arquitectura de una red neuronal convolucional es la organización de las capas. Por ejemplo, en la Figura 2.5 se presenta una red neuronal convolucional con 3 capas de convolución, dos capas de pooling y al final dos capas densas.

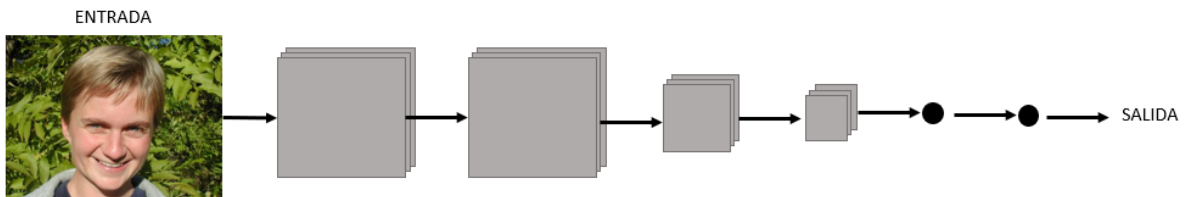


Figura 2.5 Ejemplo de arquitectura de una red neuronal convolucional

2.4 Redes Neuronales de Arquitectura Codificador - Decodificador

Las redes neuronales de arquitectura codificador - decodificador tienen por objetivo tomar la estructura original de los datos y comprimirla en una primera etapa, la codificación, y después reconstruirla volviendo a la estructura original, etapa de decodificación (Olshausen y Field 1996, Lee et al. 2006, Vicent et al. 2008, Bengio et al. 2014). La Figura 2.6 muestra un ejemplo de este tipo de arquitectura donde los bloques azules representan la codificación donde las imágenes reducen su tamaño y los bloques verdes representan la decodificación, donde las imágenes aumentan su tamaño hasta llegar al tamaño original de los datos. Ronneberger et al. (2015) propuso la arquitectura de red denominada U-NET, donde además de tener la etapa de codificador-decodificador se agregan puentes de conexión entre los bloques que reducen y amplían los tamaños de las imágenes (ver Figura 2.6).

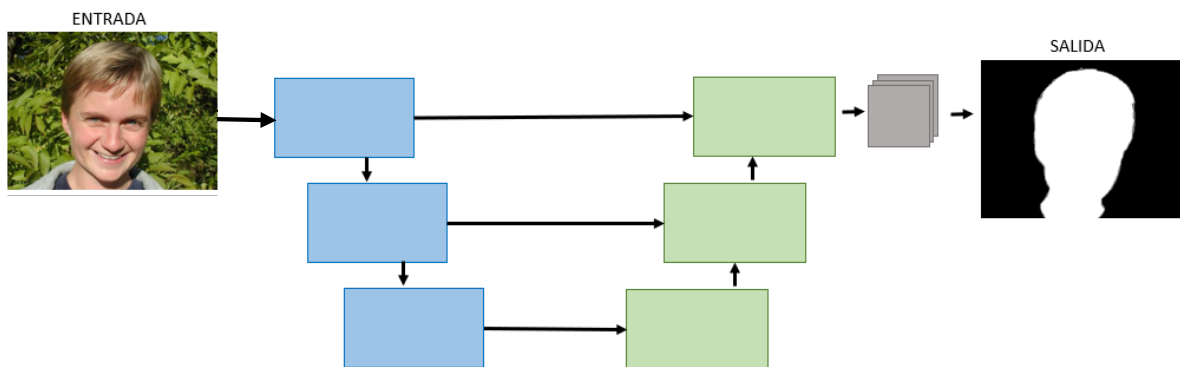


Figura 2.6 Ejemplo de red neuronal codificador-decodificador

Algunas de las aplicaciones más importantes de este tipo de arquitectura de red neuronal es la segmentación de imágenes y la eliminación de ruido. Según Sánchez et al. (2017), la segmentación de imágenes consiste en dividir una imagen en regiones con características similares, como color, textura u orientación. Más específicamente, es el proceso de etiquetar cada pixel en la imagen con el objetivo de que pixeles con etiquetas iguales tengan características similares.

Algunas de las capas que se agregan en este tipo de arquitectura son:

Capa UpSampling: El proceso de UpSampling consiste en aumentar el tamaño de una imagen. La imagen se amplía de acuerdo a cierto valor que es definido por número de renglones y columnas, o por porcentaje de aumento. La Figura 2.7 nos muestra un ejemplo donde el tamaño de la imagen se duplica, por lo que cada píxel en la imagen original se representa 4 veces en la imagen nueva. A este súper muestreo que replica el valor del pixel en tantos pixeles sea necesario para lograr la ampliación deseada se le denomina interpolación mediante vecino más cercano.

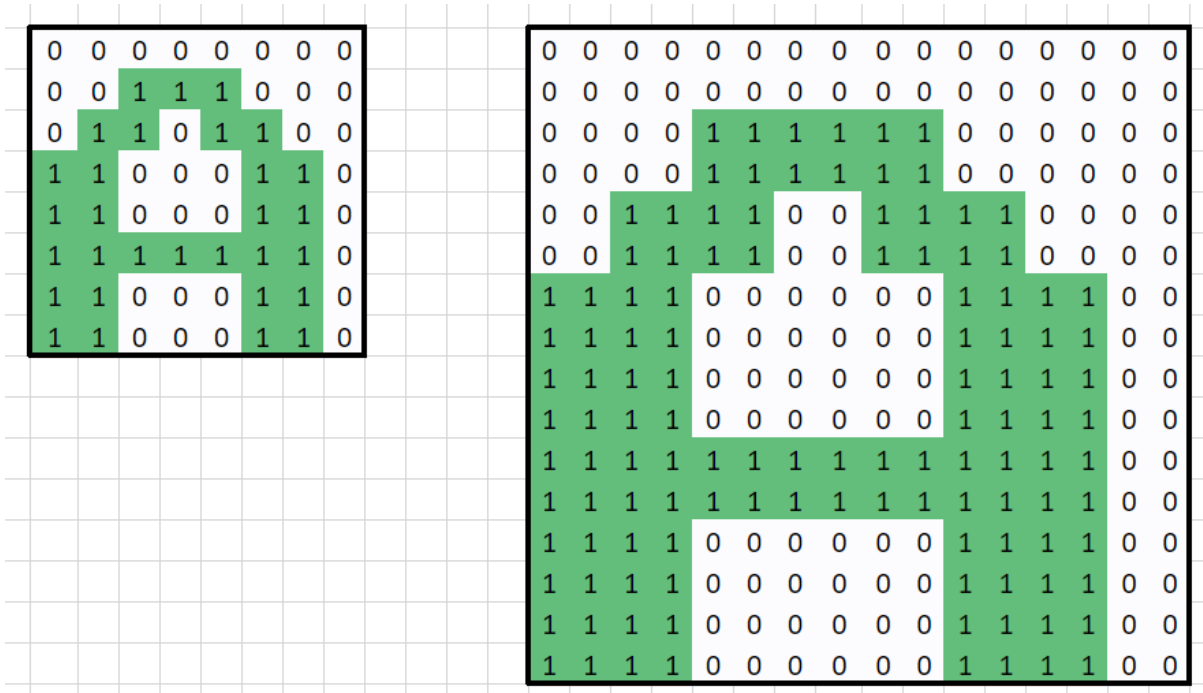


Figura 2.7 Ejemplo de UpSampling

Capa Concatenación: Esta capa tiene por objetivo manipular diferentes arreglos de tal forma que queden unidos en uno sólo. La Figura 2.8 muestra un ejemplo de esto.

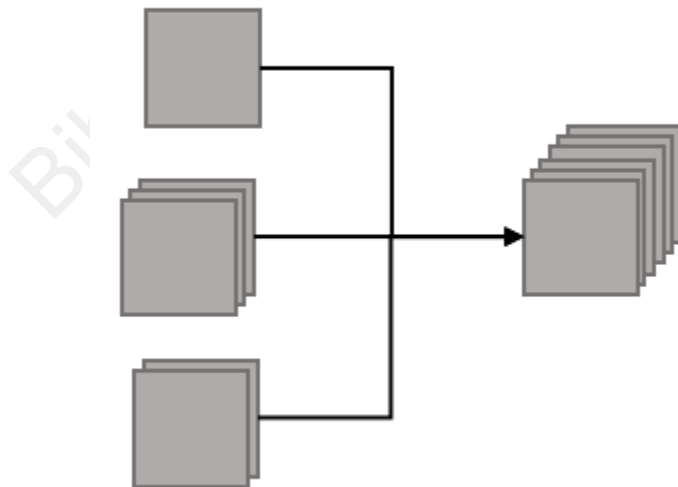


Figura 2.8 Ejemplo de Concatenación

2.5 Estado del Arte

En esta sección se presentan los documentos que presentan avances significativos en la identificación y segmentación de neumotórax en imágenes médicas. Identificación y segmentación son problemas muy distintos. Por un lado, la identificación es cuando se determina por medio de un algoritmo si la imagen contiene o no neumotórax sin entrar en detalle de la zona afectada. Por otro lado, la segmentación consiste en diferenciar la zona afectada por neumotórax píxel por píxel en la imagen. Este último es el objetivo de esta investigación.

Lindsey et al. (2018) proponen una metodología que utiliza aprendizaje profundo para determinar si una imagen contiene o no neumotórax. Utilizan información de ultrasonido torácico para alimentar su algoritmo. Sus resultados son 98.3% de precisión utilizando imágenes y 99.8% utilizando vídeos. Thrall et al. (2019) desarrollaron una red convolucional profunda para determinar si una imagen contenía o no neumotórax. Ellos utilizaron imágenes de tomografía computarizada para entrenar el algoritmo, cada imagen de 36x36 píxeles. Sus resultados fueron 100% de sensibilidad (160 de 160 neumotórax detectados) y especificidad de 82.5% (33 verdaderos negativos de 40). Más recientemente, Thian et al. (2022) propusieron nuevamente una red convolucional profunda para la clasificación de imágenes de Rayos X para determinar si se tiene o no neumotórax. Sus resultados fueron muy favorables, obteniendo valores en la curva AUC arriba de 86%.

A continuación, se presentan los trabajos que se han realizado con el objetivo de realizar la segmentación de neumotórax, es decir, determinar píxel por píxel la zona afectada por la enfermedad. Ouyang et al (2019) propusieron una red neuronal profunda para determinar las zonas en las imágenes con neumotórax. Sus resultados son 0.64 utilizando el coeficiente IoU (también conocido como Jaccard). Con el objetivo de desarrollar rápidamente metodologías para resolver el problema de segmentación de neumotórax, la SIIM (por sus siglas en inglés Society for Imaging Informatics in Medicine) lanzó en 2019 una competencia a través de la página Kaggle denominada SIIM-ACR Pneumothorax Segmentation. En seguida se presentan las publicaciones más sobresalientes de los resultados de dicha competencia. Tolkachev et al. (2020) presentaron una red neuronal profunda con arquitectura U-NET obteniendo resultados de 0.63 en el coeficiente DICE. Sus resultados quedaron en el lugar 16vo de la competencia. Por otro lado, ahora mencionamos algunos documentos que reportan los mejores resultados en la competencia. Abedalla et al. (2021) reportan un valor en coeficiente DICE de 0.8608 quedando en el top 1% de la competencia. Ellos utilizan un ensamble denominado Ens4B-UNET que incluye cuatro modelos de segmentación codificador-decodificador. Todos sus modelos fueron pre-entrenados con la base de datos de ImageNet. Además de incluir preprocesamiento como aumentación de datos. Wang et al. (2020) obtuvieron el 2do lugar con un coeficiente DICE de 0.8883. Ellos utilizaron un modelo que consiste en cuatro U-NETS y un Deeplabv3+.

Nuestra propuesta, utiliza también la base de datos de la competencia SIIM-ACR Pneumothorax Segmentation, obteniendo resultados similares a Tolkachev et al. (2020) con el 16vo lugar. Al igual que

ellos, nuestro objetivo es utilizar sólo un modelo de red neuronal cuya arquitectura es codificador-decodificador. Como se puede observar en el párrafo anterior, los mejores resultados de la competencia se obtienen por medio de ensambles de redes neuronales profundas. Desde nuestro punto de vista, una red neuronal profunda es un modelo muy complejo, y si combinamos varios modelos del mismo tipo, se obtiene un modelo que es cada vez más complejo y que requiere más tiempo de entrenamiento. Es por esto que nuestro objetivo es trabajar con un sólo modelo que dé buenos resultados.

Biblioteca Aguascalientes

Capítulo 3

Metodología

Biblioteca de Aguascalientes

3. Metodología

En este Capítulo se explica a detalle la propuesta para resolver el problema de identificación de la zona de neumotórax en imágenes de Rayos X utilizando Aprendizaje Profundo con una arquitectura codificador - decodificador. Comenzamos explicando el conjunto de imágenes utilizado y la forma en como estas imágenes se separaron en conjuntos para entrenamiento, prueba y validación. Y finalmente se describe a detalle la estructura de red neuronal utilizada para resolver el problema de segmentación de neumotórax.

3.1 Conjunto de Imágenes

Los datos se tomaron de Kaggle, del concurso SIIM-ACR Pneumothorax Segmentation propuesto por la SIIM (Society for Imaging Informatics in Medicine) cuyo objetivo es desarrollar un sistema de Inteligencia Artificial capaz de segmentar automáticamente la zona de neumotórax en imágenes de rayos X. Contiene 12,089 imágenes de rayos X de la zona pulmonar. De las cuales 2669 tenían neumotórax y 9,420 no lo tenían. Las imágenes están almacenadas en formato DICOM (Digital Imaging and Communication In Medicine) cuya extensión es dcm. Este es un estándar de transmisión de imágenes médicas y datos entre hardware médico. Contiene una imagen del paciente, y algunos datos relevantes a este como pueden ser edad, género, etc. En nuestro caso, sólo utilizamos las imágenes de rayos X cuyo tamaño original es de 1024 x 1024 píxeles. Las imágenes dcm se convirtieron en jpg para poder procesarlas más fácilmente y facilitar la manipulación de las imágenes. Para los experimentos, se redujo el tamaño de las imágenes a 256 x 256.

A la vez contienen una máscara que identifica la ubicación del neumotórax en la imagen, si es que esta tiene. La máscara está comprimida en formato Run-length encoding (RLE por sus siglas en inglés), que es una forma muy simple de compresión de secuencias de datos con el mismo valor consecutivo, lo que lo hace perfecto para comprimir una imagen binaria para las máscaras. En este caso, la representación de la imagen como una secuencia de números se ilustra en la Figura 4.1 y se describe a continuación. El primer número indica la cantidad de píxeles negros consecutivos que hay. El segundo, la cantidad de píxeles blancos consecutivos. Y así sucesivamente hasta llegar al final.

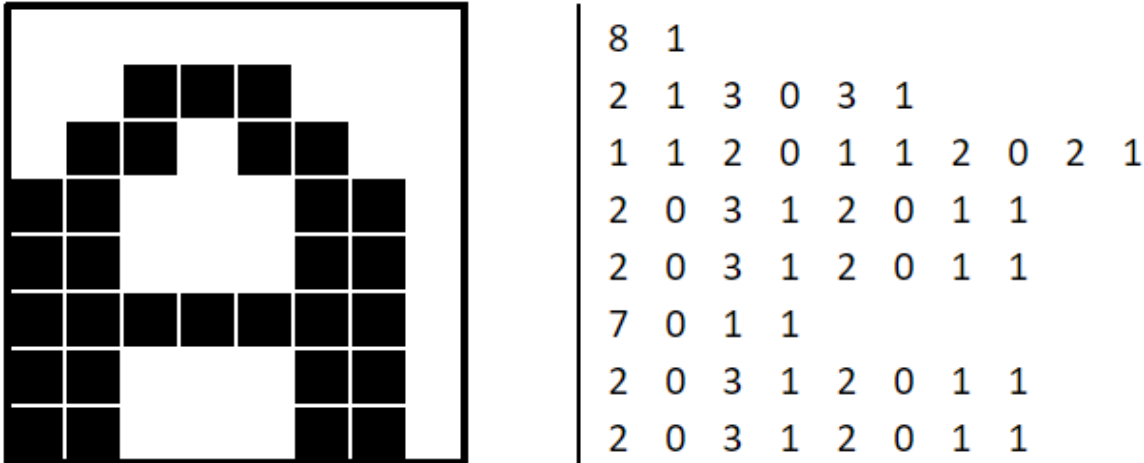


Figura 3.1 Ejemplo de formato de representación RLE. Imagen propia

3.2 Conjunto de datos para entrenar la red neuronal para segmentación

El experimento consiste en generar una red neuronal de segmentación por lo que se decidió quedarse sólo con 2669 imágenes que sí contenían neumotórax dividiendo aleatoriamente el conjunto de imágenes en:

- 1601 imágenes para el conjunto de entrenamiento.
- 534 imágenes para el conjunto de prueba.
- 534 imágenes para el conjunto de validación.

3.3 Segmentación de Imágenes de Neumotórax

Según Sánchez et al. (2017), la segmentación de imágenes consiste en dividir una imagen en regiones con características similares, como color, textura u orientación. Más específicamente, es el proceso de etiquetar cada píxel en la imagen con el objetivo de que píxeles con etiquetas iguales tengan características similares.

Específicamente, en el caso de las imágenes de rayos X, el objetivo es determinar los píxeles que corresponden a la zona de neumotórax y los píxeles que no corresponden a la zona de neumotórax (ver Figura 3.1).

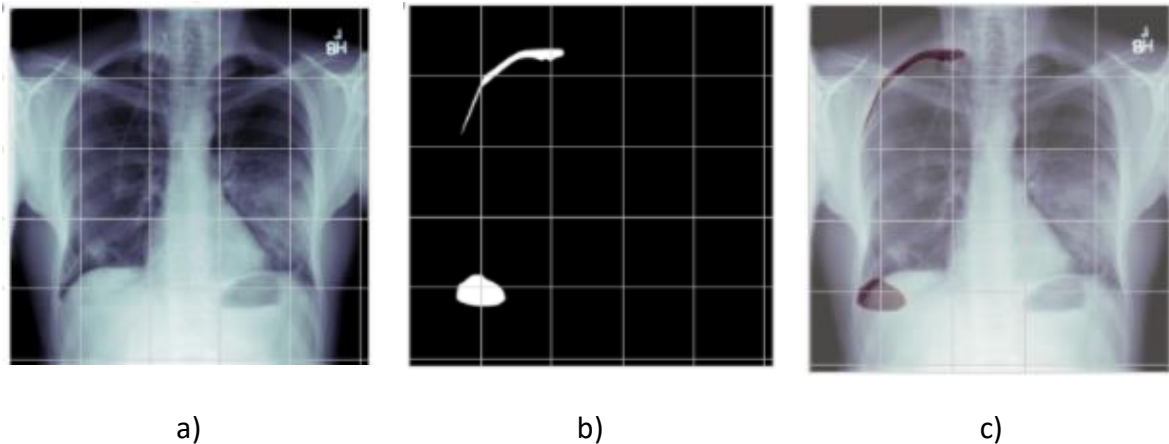


Figura 3.2 Imagen de Rayos X con neumotórax.

- a) Imagen original en rayos X. b) Imagen que representa la máscara indicando la zona del neumotórax. c) Combinación de a) y b).

3.4 Arquitectura de Red Neuronal Codificador - Decodificador

Para detectar automáticamente la zona de neumotórax en imágenes de Rayos X, proponemos una arquitectura de red neuronal tipo U-NET. El modelo utilizado es la arquitectura de la red neuronal propuesta por Reyes-Figueroa et al. (2021). En total, la red neuronal está conformada por 783,986 parámetros. Y su arquitectura se puede observar en las imágenes 3.2, 3.3, 3.4 y 3.5, la cual será detallada a continuación.

En la Figura 3.2 se representa la arquitectura de red neuronal codificador-decodificador utilizada. En la Figura 3.3 se muestran a detalle los bloques de las capas de la red neuronal. En ambas figuras (3.2 y 3.3), los bloques azules representan la codificación inicial de la imagen, reduciendo su tamaño y procesando los datos con una combinación de capas de convolución, dropout, pooling, entre otras. Cada bloque reduce el tamaño de las imágenes a la mitad. En el primer bloque las imágenes tienen un tamaño de 256x256 píxeles y al final del bloque 5 llegan a un tamaño de 8x8 píxeles. La Figura 3.4 muestra a detalle las capas y parámetros de las primeras capas de la arquitectura de la red, así como el contenido de los bloques azules. Cabe mencionar que los bloques azules son todos similares en cuanto a capas y tamaño de convoluciones, con la única diferencia del tamaño de las imágenes.

Por otro lado, los bloques verdes (ver Figuras 3.2, 3.3 y 3.5) representan la decodificación de la imagen, volviéndolo a su tamaño original usando capas similares a la codificación, usando UpSampling en lugar de pooling, concatenando la imagen de la codificación para no perder la información original. La Figura 3.5 muestra el detalle de cada uno de los bloques verdes y también las capas finales de la red neuronal. Nuestra propuesta consiste en entrenar la red neuronal desde 0 utilizando imágenes de rayos X de Neumotórax.

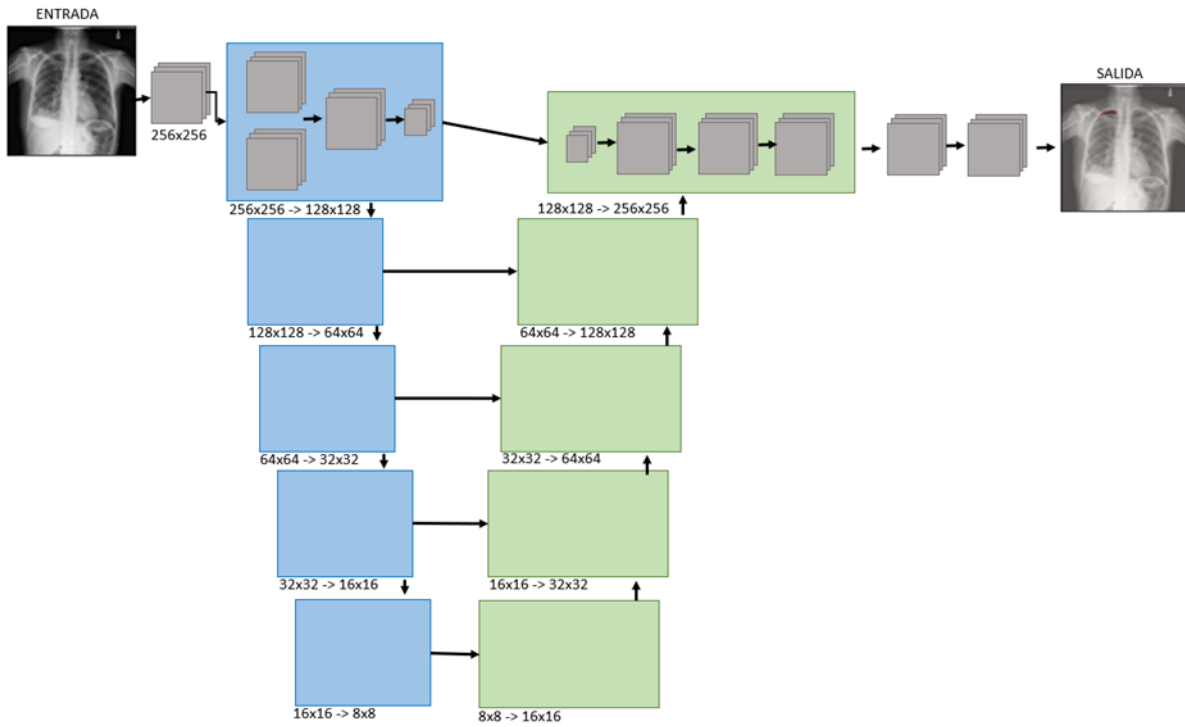


Figura 3.3 Arquitectura de la red neuronal utilizada para la segmentación de Neumotórax

Biblioteca Agila

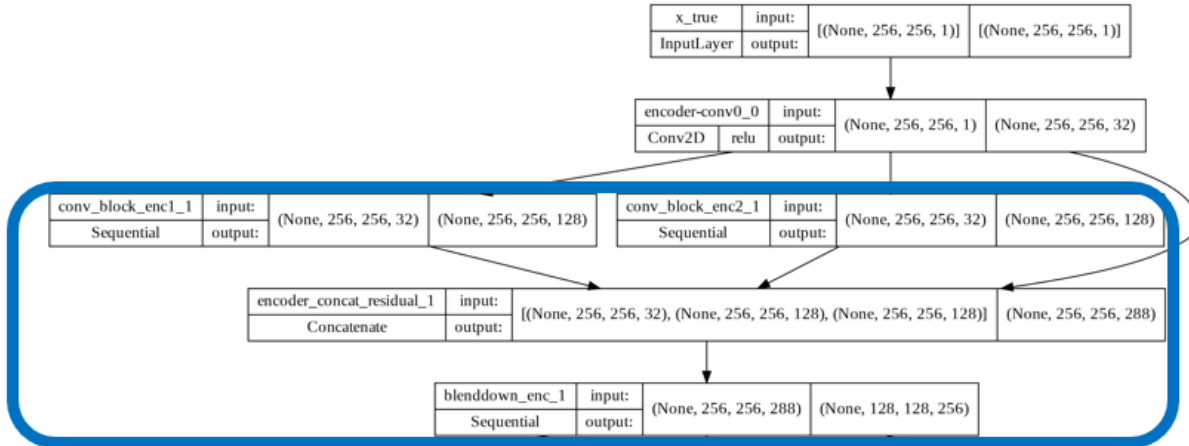


Figura 3.5 Inicio de la arquitectura de la red neuronal utilizada para la segmentación de Neumotórax con detalle

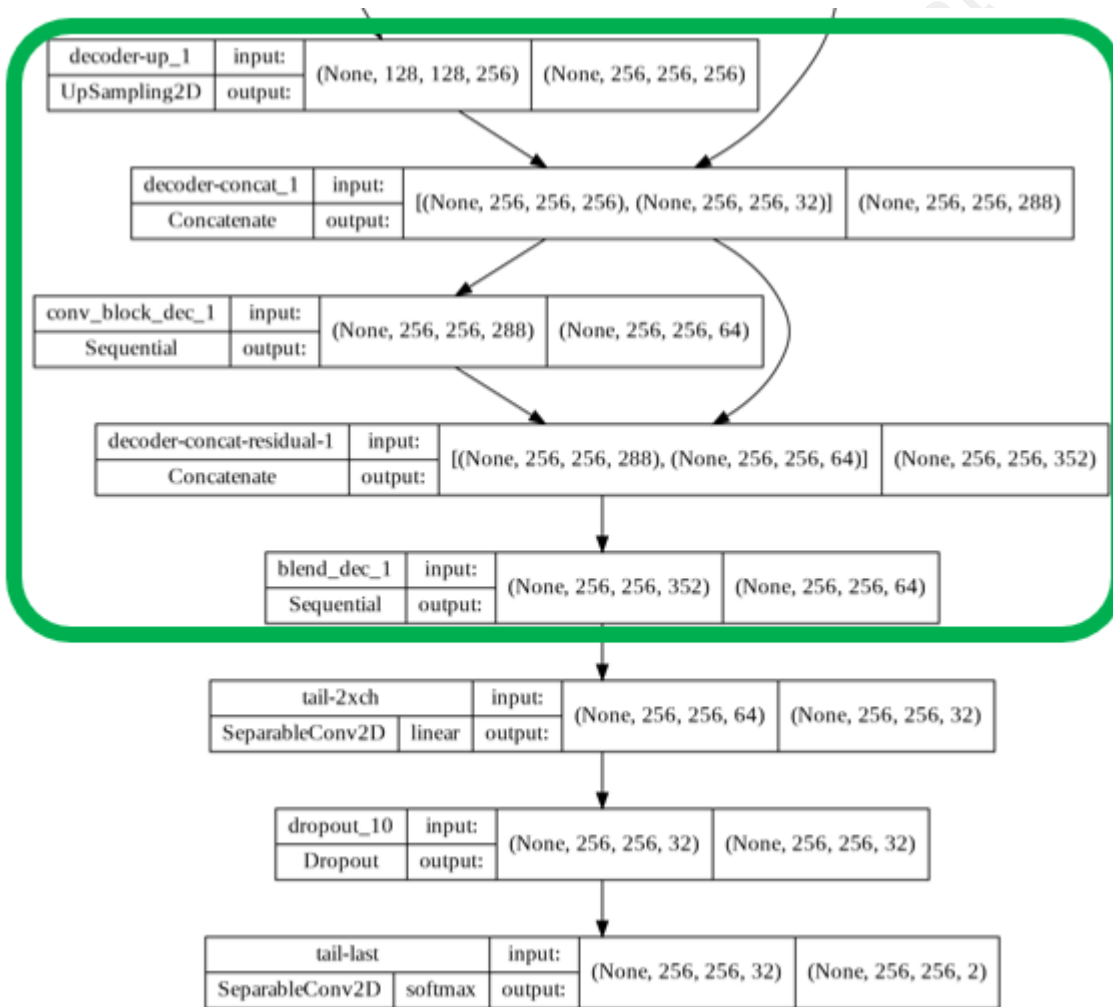


Figura 3.6 Fin de la arquitectura de red neuronal utilizada para la segmentación de Neumotórax con detalle

Capítulo 4

Experimentos y Resultados

4. Experimentos y Resultados

En este capítulo se presentan los resultados de nuestro modelo de segmentación de neumotórax en imágenes de rayos X. Primero se describen las características del equipo de cómputo utilizado para el experimento. Luego se presentan los resultados del modelo iniciando con la convergencia. Después se realiza la descripción del coeficiente de Jaccard utilizado para medir los resultados de segmentación. Y por último se muestran los resultados utilizando la métrica Jaccard para los diferentes conjuntos de imágenes junto con un listado de imágenes donde se presenta la segmentación realizada por nuestro modelo.

4.1 Equipo de cómputo

- Procesador: Intel(R) Core (TM) i9-9960 X CPU @ 3.10GHz
- Memoria RAM: 128 GB
- Sistema operativo: Windows 10 Pro.

4.2 Convergencia de la red neuronal

Para entrenar la red neuronal se utilizó el algoritmo de optimización RMSPROP utilizando como función objetivo de pérdida (loss) Sparse Categorical Crossentropy. La Figura 4.1 muestra la convergencia de la optimización tanto en el conjunto de entrenamiento como en el conjunto de validación. En total se ejecutaron 60 épocas. Se puede ver que desde aproximadamente la época 20 ya se tiene convergencia en el conjunto de validación.

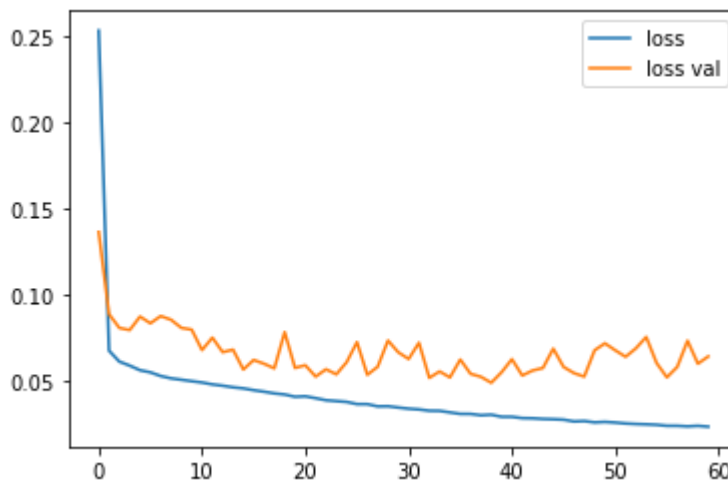


Figura 4.1 Convergencia del entrenamiento de la red neuronal

4.3 Resultados en términos del coeficiente Jaccard

La métrica de evaluación de nuestros experimentos es el coeficiente de similitud Jaccard. El cuál está definido en la Ecuación 3.1, donde las operaciones intersección y unión entre conjuntos está representado en la Figura 4.2. Los valores del coeficiente Jaccard oscilan entre 0 y 1, siendo 0 el peor resultado y 1 el mejor.

$$J(A, B) = \frac{A \cap B}{A \cup B} \quad (5.1)$$

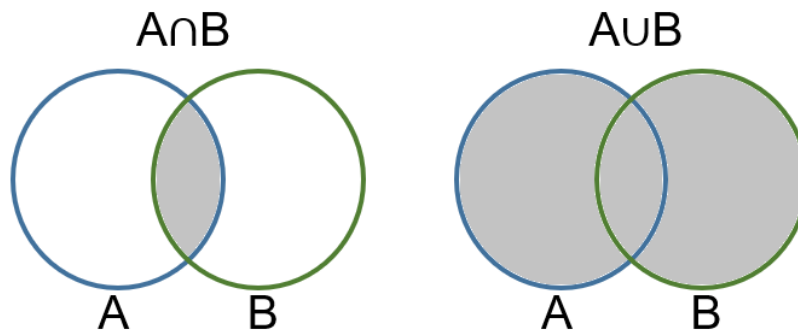


Figura 4.2 Operaciones de intersección y unión en conjuntos.

En el problema de segmentación de neumotórax, se pueden calcular dos valores del coeficiente de Jaccard por imagen. Uno para la zona de píxeles con neumotórax y otro para la zona de píxeles sin neumotórax. En el primer caso, para los píxeles con neumotórax, el conjunto A y el conjunto B de la Ecuación 5.1 representan los píxeles que en la imagen original son zonas con neumotórax y los píxeles en la imagen resultante marcados como zonas de neumotórax. En este caso, la unión sería la suma de todos los píxeles que están marcados como neumotórax en una u otra imagen. Y la intersección sería la suma de píxeles que están marcados como zona de neumotórax en ambas imágenes. Cuando se tiene más de una zona de interés, como en este caso, los resultados se pueden expresar utilizando un promedio denominado macro. En este sentido, el coeficiente Macro-Jaccard es el resultado del promedio de todas las clases. Específicamente en este problema, sería el promedio de los resultados de la segmentación de la zona con neumotórax y la zona sin neumotórax.

La Figura 4.3 muestra los resultados de nuestros experimentos expresados en el coeficiente Macro-Jaccard para los conjuntos de entrenamiento, validación y prueba. Así mismo, la Tabla 5.1 muestra los promedios de dichos resultados. Se puede ver que como era de esperarse, los resultados del conjunto de entrenamiento son mucho mayores que los resultados de los conjuntos de validación y prueba, ya que en este conjunto se encuentran las imágenes que se utilizaron para entrenar los parámetros de la red neuronal.

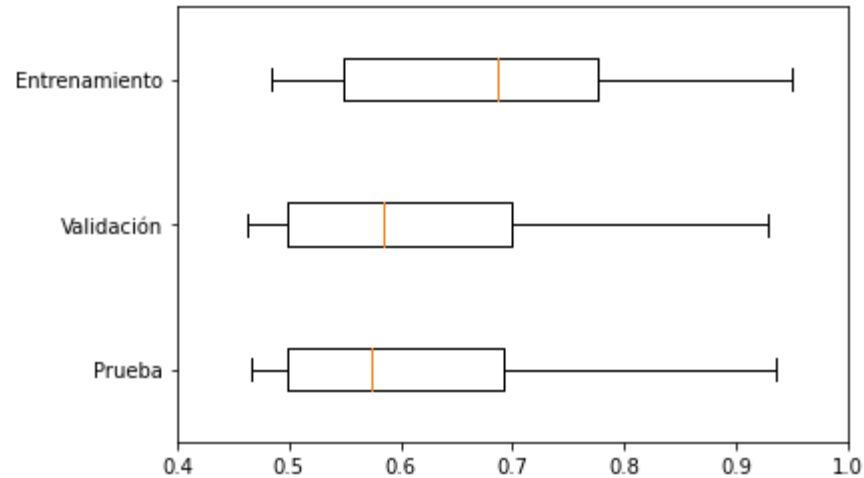


Figura 4.3 Macro-Jaccard de los conjuntos de imágenes de entrenamiento, validación y prueba.

Tabla 4.1 Promedio de Macro-Jaccard de los conjuntos de imágenes de entrenamiento, validación y prueba.

Conjunto de datos	Coeficiente Macro-Jaccard
Entrenamiento	0.6755
Validación	0.6108
Prueba	0.6072

4.4 Resultados en imágenes

En esta sección se muestran los resultados de la segmentación del neumotórax en imágenes. La Figura 4.4 muestra varios ejemplos de la segmentación utilizando nuestra propuesta. A la izquierda, se muestra la imagen original de Rayos X. En el centro se puede ver la zona de neumotórax real resaltada con una máscara en rojo. Finalmente, a la derecha, se muestra la segmentación calculada con nuestra red neuronal propuesta, también utilizando una máscara en rojo para resaltar la zona.



Bibli





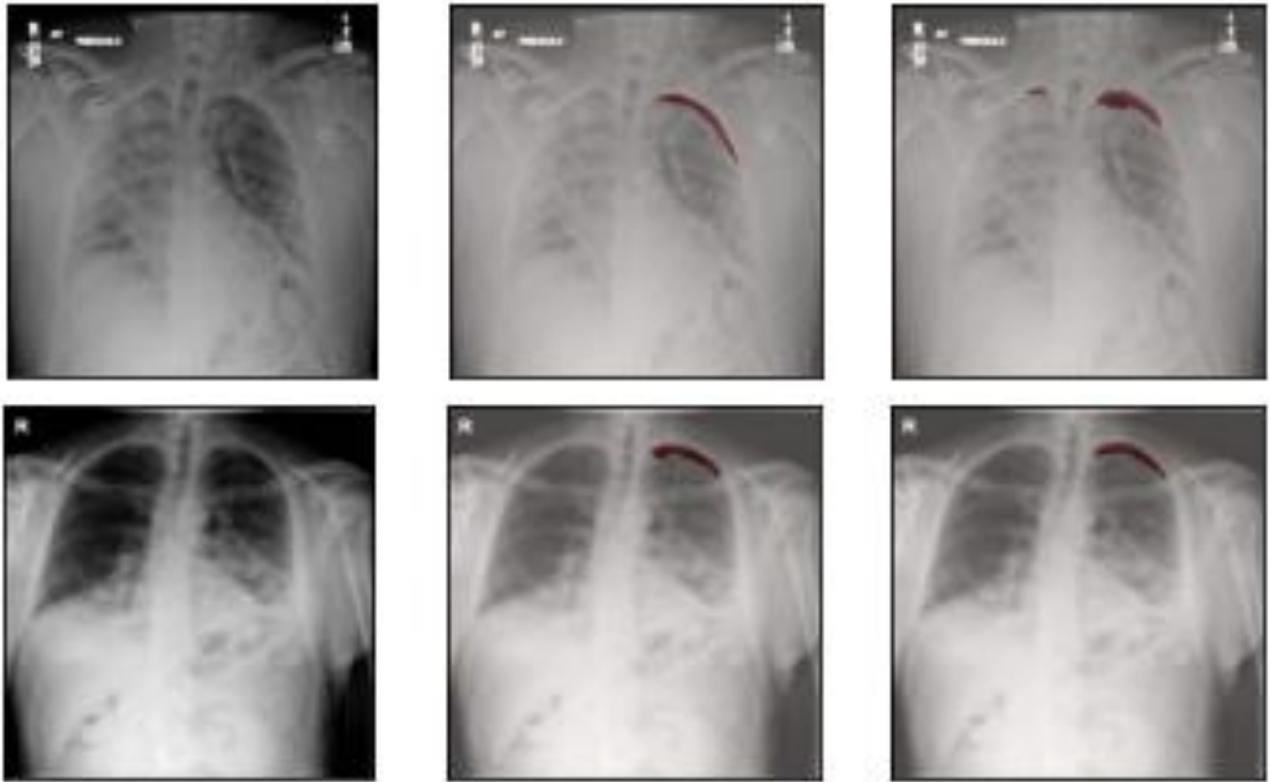


Figura 4.4 Resultados de la segmentación de neumotórax.

A la izquierda se muestra la imagen original en rayos X. Al centro se muestra la zona de neumotórax real. A la derecha se muestra la zona de neumotórax detectada por el modelo propuesto.

Capítulo 5

Conclusión

Biblioteca Vascales

5. Conclusión

Para finalizar este documento, a continuación, se enlistan nuestras conclusiones y el trabajo a futuro.

5.1 Conclusión

- Las redes neuronales codificador-decodificador son buenas y rápidas para analizar imágenes médicas.
- Esta tecnología se puede extender a otras aplicaciones médicas que tienen por objetivo, detectar zonas afectadas por enfermedades, siempre y cuando se cuente con una cantidad importante de datos previamente etiquetados.
- El problema de la segmentación automática de neumotórax es bastante retador, sin embargo, el presente documento demuestra que es posible resolverlo utilizando redes neuronales profundas.

5.2 Trabajo a futuro

- Como se mostró en la sección 1.1, existen modelos mucho más robustos para resolver este tipo de problemas usando una combinación de varios modelos codificador-decodificador para obtener mejores resultados, por un coste computacional mucho más elevado.
- Nuestro objetivo es analizar los resultados de la segmentación para mejorar la métrica utilizada en la función objetivo del algoritmo de entrenamiento de la red neuronal. La primera idea consiste en modificar los pesos de cada una de las clases de tal forma que la zona de neumotórax tenga un peso mayor para el cálculo de la métrica. Esta idea surge porque las imágenes calculadas presentan una segmentación más pequeña que lo que presenta la imagen original.

Bibliografía

Biblioteca Aguascalientes

6. Bibliografía

Abedalla, A., Abdullah, M., Al-Ayyoub, M., & Benkhelifa, E. (2021). Chest X-ray pneumothorax segmentation using U-Net with EfficientNet and ResNet architectures. *PeerJ Computer Science*, 7, e607.

Alazab, M., Awajan, A., Mesleh, A., Abraham, A., Jatana, V., & Alhyari, S. (2020). COVID-19 prediction and detection using deep learning. *International Journal of Computer Information Systems and Industrial Management Applications*, 12(June), 168-181.

Altinkaya, E., Polat, K., & Barakli, B. (2020). Detection of Alzheimer's disease and dementia states based on deep learning from MRI images: a comprehensive review. *Journal of the Institute of Electronics and Computer*, 1(1), 39-53.

Antipov, G., Baccouche, M., & Dugelay, J. L. (2017, September). Face aging with conditional generative adversarial networks. In *2017 IEEE international conference on image processing (ICIP)* (pp. 2089-2093). IEEE.

Balaban, S. (2015). Deep learning and face recognition: the state of the art. *Biometric and surveillance technology for human and activity identification XII*, 9457, 68-75.

Bengio, Y., Laufer, E., Alain, G., & Yosinski, J. (2014, June). Deep generative stochastic networks trainable by backprop. In *International Conference on Machine Learning* (pp. 226-234). PMLR.

Brinker, T. J., Hekler, A., Utikal, J. S., Grabe, N., Schadendorf, D., Klode, J., ... & Von Kalle, C. (2018). Skin cancer classification using convolutional neural networks: systematic review. *Journal of medical Internet research*, 20(10), e13936.

Elton, D. C., Boukouvalas, Z., Fuge, M. D., & Chung, P. W. (2019). Deep learning for molecular design—a review of the state of the art. *Molecular Systems Design & Engineering*, 4(4), 828-849.

Grigorescu, S., Trasnea, B., Cocias, T., & Macesanu, G. (2020). A survey of deep learning techniques for autonomous driving. *Journal of Field Robotics*, 37(3), 362-386.

Gurney, K. (2018). *An introduction to neural networks*. CRC press.

Işın, A., Direkoğlu, C., & Şah, M. (2016). Review of MRI-based brain tumor image segmentation using deep learning methods. *Procedia Computer Science*, 102, 317-324.

Lee, H., Battle, A., Raina, R., & Ng, A. (2006). Efficient sparse coding algorithms. *Advances in neural information processing systems*, 19.

Li, X., Thrall, J. H., Digumarthy, S. R., Kalra, M. K., Pandharipande, P. V., Zhang, B., ... & Li, Q. (2019). Deep learning-enabled system for rapid pneumothorax screening on chest CT. *European journal of radiology*, 120, 108692.

Lindsey, T., Lee, R., Grisell, R., Vega, S., & Veazey, S. (2018, November). Automated pneumothorax diagnosis using deep neural networks. In *Iberoamerican Congress on Pattern Recognition* (pp. 723-731). Springer, Cham.

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Rafal Jozefowicz, Yangqing Jia, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Mike Schuster, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. *TensorFlow: Large-scale machine learning on heterogeneous systems*, 2015. Software available from [tensorflow.org](https://www.tensorflow.org).

McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4), 115-133.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381(6583), 607-609.

Ouyang, X., Xue, Z., Zhan, Y., Zhou, X. S., Wang, Q., Zhou, Y., ... & Cheng, J. Z. (2019, October). Weakly supervised segmentation framework with uncertainty: A study on pneumothorax segmentation in chest x-ray. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 613-621). Springer, Cham.

Reyes-Figueroa, A., Flores, V. H., & Rivera, M. (2021). Deep neural network for fringe pattern filtering and normalization. *Applied Optics*, 60(7), 2022-2036.

Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.

Sánchez, C. N., Rivera, M., & Velázquez, R. (2017, November). Robust multiband image segmentation method based on user clues. In *2017 IEEE 37th Central America and Panama Convention (CONCAPAN XXXVII)* (pp. 1-6). IEEE.

Socher, R., Huval, B., Bath, B., Manning, C. D., & Ng, A. (2012). Convolutional-recursive deep learning for 3d object classification. *Advances in neural information processing systems*, 25.

Society for Imaging Informatics in Medicine (SIIM) (2019). SIIM-ACR Pneumothorax Segmentation. Kaggle. <https://www.kaggle.com/c/siim-acr-pneumothorax-segmentation>

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1), 1929-1958.

Thian, Y. L., Ng, D. W., Hallinan, J. T. P. D., Jagmohan, P., Sia, S. Y., Mohamed, J. S. A., ... & Feng, M. (2022). Effect of Training Data Volume on Performance of Convolutional Neural Network Pneumothorax Classifiers. *Journal of Digital Imaging*, 1-12.

Tolkachev, A., Sirazitdinov, I., Kholiavchenko, M., Mustafaev, T., & Ibragimov, B. (2020). Deep learning for diagnosis and segmentation of pneumothorax: the results on the kaggle competition and validation against radiologists. *IEEE Journal of Biomedical and Health Informatics*, 25(5), 1660-1672.

Vincent, P., Larochelle, H., Bengio, Y., & Manzagol, P. A. (2008, July). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th international conference on Machine learning* (pp. 1096-1103).

Wang, X. (2016). Deep learning in object recognition, detection, and segmentation. *Foundations and Trends in Signal Processing*, 8(4), 217-382.

Wang, X., Yang, S., Lan, J., Fang, Y., He, J., Wang, M., ... & Han, X. (2020). Automatic segmentation of pneumothorax in chest radiographs based on a two-stage deep learning method. *IEEE Transactions on Cognitive and Developmental Systems*.

Xin, Y., Kong, L., Liu, Z., Chen, Y., Li, Y., Zhu, H., ... & Wang, C. (2018). Machine learning and deep learning methods for cybersecurity. *IEEE Access*, 6, 35365-35381.